GROUP THEORY

An example based approach

$\circ$	(1)	(12)(34)	(13)(24)	(14)(23)
(1)	(1)	(12)(34)	(13)(24)	(14)(23)
(12)(34)	(12)(34)	(1)	(14)(23)	(13)(24)
(13)(24)	(13)(24)	(14)(23)	(1)	(12)(34)
(14)(23)	(14)(23)	(13)(24)	(12)(34)	(1)

Mark O'Brien

June 4, 2012

1 Introduction

I have been asked many times why people who teach math use color as much as they do during their lectures. Why is it that then that upper level undergraduate math books make almost no use of color?

These notes were written for my math 330 class which I teach at the University of Illinois at Chicago. They focus on developing non-trivial examples to accompany the usual theorems that one comes across in a first course in group theory. While a graduate student will not learn much from the theory portions of the text, even they should learn quite a bit from the examples presented. I have been writing this text because most undergraduate texts develop only 2 or 3 non-trivial examples of non-abelian finite groups: the symmetric, dihedral, and alternating groups. At the same time, this leads to almost no examples of any non-trivial homomorphisms or quotient groups. This text also deals with more general subgroups of the symmetric groups (given in terms of generators), various quaternion groups (binary groups), as well as matrix groups with real, complex, rational, integer, and modular integers coefficients.

I break nearly every writing rule in the book. I make extensive use of contractions, I use a ton a footnotes (92 at present ... eat that AMS), and I even use the word ass once ... ok, now twice. I put commas (and not periods to spite South Americans and main-land Europeans) into really huge numbers at least 4,231,552 times, and I use the words "really" and "very" a lot. I have tried to not make up any words, but I assume I have failed at that ☹. Oh, and I use several smiley and sad faces as well. I talk about left brain versus right brain problems. I have tried to keep the language loose while, at the same time, keeping the mathematics tight. This writing style is based on the physics textbooks by David Griffiths whose combination of formal and informal writing is something that is lacking in undergraduate math textbooks, especially in group theory.

Basic Sets and Notation

The most basic sets in this course are the following:

(1) The natural numbers, $\mathbb{Z}_n \stackrel{\text{def}}{=} \{0, 1, 2, 3, \ldots\}$;[I]

(2) The integers, $\mathbb{Z} \stackrel{\text{def}}{=} \{0, \pm 1, \pm 2, \pm 3, \ldots\}$;

(3) The rationals, $\mathbb{Q} \stackrel{\text{def}}{=} \{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\}$;

(4) The reals, $\mathbb{R}$;

For all $a \leq b \in \mathbb{R} \cup \{-\infty, \infty\}$ we also define

[I] It seems every person is different as to whether or not they include 0 in $\mathbb{Z}_n$

(5) The intervals

$$[a,b] \stackrel{\text{def}}{=} \{x \in \mathbb{R} : \ a \le x \le b\} \qquad [a,b) \stackrel{\text{def}}{=} \{x \in \mathbb{R} : a \le x < b\}$$
$$(a,b) \stackrel{\text{def}}{=} \{x \in \mathbb{R} : a < x < b\} \qquad (a,b] \stackrel{\text{def}}{=} \{x \in \mathbb{R} : a < x \le b\};$$

(6) The integer interval $\overline{[a,b]} \stackrel{\text{def}}{=} [a,b] \cap \mathbb{Z}$;

(7) For any integer $n > 0$, $\overline{n} \stackrel{\text{def}}{=} \overline{[1,n]}$;

(8) The complexes, $\mathbb{C}$;

(9) The quaternions, $\mathbb{H}$;

(10) and $\forall n \in \mathbb{Z}$, the integers mod n, $\mathbb{Z}_n$. [II]

You are expected to be familiar with all of these sets with the exception of the last two, $\mathbb{H}$ and $\mathbb{Z}_n$, listed here ($\mathbb{Z}_n$ is really a different set for each positive integer n). Note that saying that $m \in \overline{n}$ is exactly the same thing as saying that m is an integer $1 \le m \le n$, i.e.

$$\overline{n} = \{m \in \mathbb{Z} : 1 \le m \le n\}.$$

For any set S, $S^\times$ denotes the non-zero elements of S, that is $S^\times \stackrel{\text{def}}{=} S \setminus \{0\}$. Of course we will only make use of this notation if S contained 0 in the first place. For instance, if S were the set of blades of grass on the residence of the American citizen with social security number 022 – 69 – 9424 at exactly 7PM on September 17th, 2001, it would be silly to talk about $S^\times$. When it makes sense, for instance if $S = \mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$, the notation S^+ means the set of positive elements of the indicated set. For instance $\mathbb{Z}^+$ denotes the set of positive integers. The notations $\mathbb{C}^+$, $\mathbb{H}^+$, and $\mathbb{Z}_n^+$ make no sense, and are not defined.

Note that we could have denoted $\mathbb{Z}^+$ as $\overline{\infty}$, $\mathbb{R}$ as $(-\infty, \infty)$, $\mathbb{Z}$ as $\overline{[-\infty, \infty]}$, as well as many other fun plays on the definitions.

More on Prerequisites

We assume a basic understanding of logic including the meanings of and, or, and implies. We give a quick review of set theory at the beginning, but clearly not enough to get an independent grasp of it. Essentially no other prerequisites are required. However some basic understanding of matrices and vector algebra would be helpful in order to understand some of the examples.

There are several types of exercises in this text. The exercises which require a proof or a calculation should be familiar. A third type of exercise requires you to find a mistake in a proof given. These problems highlight common mistakes students make. The goal is find the poor step which constitutes an error and explain why it is an error. The goal is **not** to explain why the theorem (conclusion) is false. Therefore the answer would go as "line (such and such) is an error since (such and such) does not hold in our current situation."

Sections marked with a * are optional on first reading. However in these sections the student is still expected to have the definitions presented "on their radar screen."

Types of Definitions

In general, in mathematics every concept has a large pool of possible definitions. Very roughly, definitions can be categorized in one of four ways

1. Intuitive This is the kind of definition that could be given as quickly as possible. Often the definition to be given to somebody with very little background. For instance, if you were going to define the derivative to somebody with only a basic understanding of high school algebra you would probably tell them its the slope of the tangent line.

[II] The letter Z in $\mathbb{Z}$ stands for *Zahlen*, the german word for "number." The Q in $\mathbb{Q}$ stands for "quotients." The letter H in $\mathbb{H}$ is in honor of William Rowan Hamilton who first described the quaternions in 1843.

2. Working These are the definitions that you can effectively use to prove things. This definition is often more abstract than the intuitive definition. For instance, the working definition of the derivative is given in terms of limits and difference quotients:

$$f'(a) := \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

While the definition of the derivatives in terms of tangent lines is "intuitive" and necessary to grasp to get any sort of understanding of calculus, most of the basic laws of derivatives – including the chain rule, product rule, and Linear rule – are easily derived from this definition.

3. Natural In research, most definitions the mathematician comes up with are natural definitions. These are definitions that are already present in terms of what has been studied. In calculus, after studying the derivative for several months, it is natural to define the integral as doing a derivative in reverse. This is particularly the case after studying the mean value theorem which shows that this process is determined up to a constant. In this course we will take a natural definition of a *normal subgroup*. However, for the most part we will *use* a working definition. Of course what is natural to one is not natural to another. Few students find the definition of a homomorphism of groups to be natural the first time they see it. However after a couple of semesters of studying algebra, looking back with hindsight, the definition becomes incredibly natural. One of the main tasks of teaching abstract algebra to undergraduates is convincing the student that this definition is indeed natural. I hope by the end of the course you are convinced that we could not possibly study group theory without studying group homomorphisms (and not just because I told you that this was the case!).

4. Axiomatic Chances are you have seen very few axiomatic definitions in your basic math courses. However, this is the way that most of the main concepts are first presented in higher level mathematics. An axiomatic definition defines an object of study as something which has certain properties. In an axiomatic definition it's not even clear that such things that you are defining even exist! If you have taken Linear Algebra before you have seen this done for a vector space. In this course our main object of study, a group, is defined axiomatically. Students actually tend to like axiomatic definitions more than they actually anticipate! For instance, very few students remember the formal definition of the derivate, but most students remember the chain product and Linear rules for derivatives. It's actually a nice surprise that these students aren't as far off as you would guess! A derivative could indeed be defined as a rule that satisfies those three properties and the additional property that the derivative of the function x is 1. Actually you could get away with a single axiom to define the derivate: for any function f

$$\frac{d}{dx} \ln(f(x)) = \frac{f'(x)}{f(x)}.$$

In my opinion, the key to getting "good" at mathematics is understanding these distinctions and for a given concept being able to move freely between them. While it s human nature to have a favorite definition, it's much more helpful to move between them. Asking a mathematician what their favorite definition of the derivative is like asking a carpenter if they prefer a hammer or a screwdriver. While the carpenter will probably be able to answer this question, and they may even use one tool over the other when either could be used; a carpenter who prefers the hammer would of course use a screwdriver for a job that involves, say, screws.

Show your work ... on second thought

Starting at a young age, you had the the sentence "show all of your work" beat into you. Now is the time to beat it back out of you. When writing up a proof, your goal is to convince your classmates (not me, and certainly not yourself!) that the question is true. In doing so, be sure to allow the reader to *see the forrest for the trees*.[III] So if I asked you to prove that

$$x^2 + 3x + 2 = 0$$

has a solution, you would simply say

[III] And not use any cliché's ... oops.

The equation $x^2 + 3x + 2 = 0$ has a solution since

$$(-1)^2 + 3 \cdot (-1) + 2 = 1 - 3 + 2 = 0.$$

I don't care how you know this. I don't care if you factored the polynomial, used the quadratic equation, used a calculator, or if you randomly guessed it. If, on the other hand, I asked you a more detailed question make sure you answer that question.

In a class on group theory the burden is on me, and not on you, to make the questions I ask precise enough for you to answer. For instance, a typical Calculus I question, like compute

$$\frac{d}{dx}\left(\frac{x^3+1}{x^3-1}\right)$$

could in this class be answered as

$$\frac{d}{dx}\left(\frac{x^3+1}{x^3-1}\right) = \frac{d}{dx}\left(\frac{x^3+1}{x^3-1}\right).$$

While if I were teaching a Calculus I class this way, I would be incredibly annoyed by this answer, in this class this answer is just as good as any other answer since it's not clear what it means to compute something! In fact it is rather difficult to phrase this question in such a way that I would be guaranteed to get "what I want." Even statements like "take the derivative" are rather vague. For instance, if I asked a student to "take the integral" of $\ln xe^{-x^2}$, the answer a computing system like `MAPLE` gives us

$$\int e^{-x^2} \ln x \, dx.$$

This is due to the fact that the question has no familiar solution.[IV] In fact this **is** the best answer!

This is something I did myself many times to my graduate professors, and it's something that's actually encouraged in higher level math courses. One instance that jumps out in my mind is a problem where the professor asked us to draw 3 examples of (something called a) covering spaces of a torus and I drew the trivial example (which is the torus) three times. This is certainly correct since the professor neither asked for *distinct* examples, nor did he ask for *non–trivial* examples. The explanation, "well I implied it" is never good enough.

For many proofs, this is not how I wrote them. This is because my goal is to teach, while your goal is to mathematically convince. When you finish your problem, your best friend in the class should be able to look at your proof and have no doubt (perhaps they need to look things up) that every step that you have done is valid.

Other Books

{Make this better}

- [Rot96].
- [Gar08].
- [DF04].
- [Hun80].
- [SSA01].

[IV] Now let's figure out if the student should add $+c$ to the end of this answer, groan.

Types of Problems and Presentations

Example/Theorem. In many places in the text I have what I call "example/theorems." These doohickey's are in place more for the reader to learn how to do proofs involving the new concepts as opposed and may not fit into the flow of the text in the sense they are not used to build up the theory at the moment (however they are often used later in the text).

False Theorem . The goal for a student dealing with exercises labeled as False Proofs is to find what step I did wrong. The content of these problems presents something which is false and then I go on to "prove" this false thing. Logic, being as sweet as it is, says that I had to have made at least one false step. These problems highlight several of the common mistakes first year abstract algebra students make.

Rant. Rants are often times things I originally labeled as remarks. Over my 15 plus years of learning and teaching I have noticed a common theme amongst the "best" mathematical and physics textbooks: they often rant about things. Books like Hatcher's *Algebraic Topology*, Misner, Thorne, and Wielder's *Gravatation*, Griffiths *Quantum Mechanics*, and Munkres *Topology* book often discuss little corners of their subjects which bring in a new light to the student. Granted, in theory, the mathematician and the mathematical textbook ought to be concise. However, you will have your entire career to deal with this aspect of math and I feel that it is impossible to feel like you are getting to the center of a subject like group theory with a textbook which is overly concise in their presentation. Parts of the text that are labeled as rants serve as a potential warning to the student.

Some Common Notation and Abbreviations

Some common abbreviations I use appear in Table 1:

Abbreviation	Meaning
TFAE	The following are equivalent
PNT	Proper, non-trivial
LHS	Left hand side
RHS	Right hand side
iff	If and only if
s.t.	Such that
OST	Orbit Stabilizer Theorem, 10.92
KIT	Kernel/Image Theorem,10.107
OPST	Orbit Stabilier p-Theorem, 10.125

Table 1.1: Common Abbreviations

As you should be familar with (from say a Calculus 2 class) the notation $\sum$ corresponds to a summation:

$$\sum_{i=0}^{n} a_i^4 = a_0^4 + a_1^4 \cdots a_n^4.$$

If you have never seen the notation before, the notation $\prod$ is to multiplication as $\sum$ is to addition. For example,

$$\prod_{i=1}^{7} a_i^3 = a_1^3 \cdot a_2^3 \cdot a_3^3 \cdot a_4^3 \cdot a_5^3.$$

Or for an example with concrete numbers,

$$\prod_{i=1}^{5} i^2 = 1 \cdot 4 \cdot 9 \cdot 16 \cdot 25 = 14,400.$$

Course Outlines

I Background Material

2 Sets, Functions, and Equivalence Relations

In this chapter we review some of the basic notions that I hope were covered in a foundations course. Perhaps I cover relations more thoroughly than you may have seen them covered in that course. This is not because I intend on using them in the context presented, but I believe they give a context which is both sufficiently abstract and proofs which are reasonably straight forward as to give you guys a great test arena to try proving things which do not find immediately obvious. From there we develop the two special examples of relations which are the most useful, namely equivalence relations and functions.

2.1 Undefined Concepts: Sets and Objects

The task of a dictionary is to define words. However, a dictionary consists of nothing but words. The poor, hapless dictionary writer cannot even get their poor foot in the door without running into definitions which depend on each other. The second they write their first word down they are already "assuming" the reader knows that word. One could flip the dictionary open to the definition of that first word, but alas the first word in *that* definition is either that same word (hence a cyclic definition) or it is also undefined. Thus it is impossible to give a "perfect" dictionary.

It turns out that the mathematician has a similar plight. We cannot possibly "start at the beginning", but we must start somewhere. We, like most mathematicians choose to start with sets, objects, and logic. (Of course these can be defined in terms of even more primitive notions, but in my (incredibly ignorant) opinion this is the job of the philosopher) This is what as known as *naive set theory*. However, unlike the dictionary writer, we never bother to lead you in circles, but we simply smile and state, "you know what sets, objects, and logic are" and then we start our journey. However, to convince you that what we are doing could lead to danger (it doesn't for us!) check out Russell's paradox, Exercise 2.1 just below.

In this section I give TMI (too much information) about sets and logic. You should really think of the extra details presented here (as opposed to the foundations course you took) are curiosities. I hope to provide a slightly more formal treatment of these foundations than you have seen before, but not nearly as forma as you would see in a logic course, which in my opinion, one should take in a philosophy course.

As you will soon discover, like nearly all concepts in mathematics, a group is a set with some extra structure. The extra structure is a rule turning pairs of elements of the set into a single element of that same set, called a *binary operation* or a *multiplication*. This is something that you should already be familiar with since addition, multiplication, subtraction are examples of this for the integers, polynomials, reals, complex numbers, and rational numbers. To drive you nuts, these functions are generically all called *multiplication* (even if it's addition!). For the set with a multiplication to form a group, the multiplication must satisfy certain rules (called axioms). Later we discuss group homomorphisms. These are functions between groups which allow us to compare them. For instance, such functions will allow us to decide whether or not we should regard two groups as being the same. Since sets and functions show up in the definitions, we review these fundamental concepts in this chapter.

The notion of an *object* is left undefined. Informally, an object is anything that we can conceive (even abstractly). Every noun is an object. This includes the *word* "run". Dogs, days, suns, molecules, the set of all molecules, the set of the set of all molecules, a pink bow, and a (perfect) circle are all objects. Every thought that any person has ever had, every thing anybody has ever done, and any song a person has ever sung are also objects.

The basic notions of logic, including "and", "or", "true", and "false", are also left undefined. A proposition is a statement which is either true of false. Two objects are said to be equal, denoted with an equal sign, =, if every proposition which is true about the first object is true for the second and vice-versa. This causes some problems. For instance, with the definition, technically, Super Man and Clark Kent are not the same person since Louis Lane thinks that Super Man can fly while she thinks that Clark Kent can not fly. Therefore, the proposition "Louis Lane thinks you can fly" it true for one and not the other.[I] Of course we completely ignore these logical debates (or we

[I] This actually implies that the statement "Louis Lane thinks you can fly" is not a proposition since it's both true and false for the same person, namely for Clark Kent. But alas, we are like the nuts dog chasing our

will go nuts!)

Roughly, a set X is a proposition, also called X, applied to all objects. For example, X could be the proposition which is true for all people over the age of 17 and is false for all other objects. If x is an object we write $X(x)$ for the proposition X applied to x. Since $X(x)$ is a proposition $X(x)$ is either true of false. For example, X applied to me is true (assuming I'm not dead by the time you read this!), but X applied to a 40 year old turtle is false since a turtle is not a person. We think of X as a collection (whatever that is) of objects for which the proposition $X(x)$ is true. For example, with our running example X is considered to be the set of people over the age of 17. If $X(x)$ is true we write $x \in X$, and we say that x is an *element* of X or x is *in* X, and we write $x \in X$.[II] On the other hand, if $X(x)$ is false we write $x \notin X$ (x is not an element of X). With our running example

$$\text{Mark O'Brien} \in X \quad \text{but} \quad \text{40 year old turtle} \notin X.$$

By definition of a proposition, for a given set X every object is either in X or it is not in X. Therefore for any set X we have that the set

$$\{x : x \in X\}$$

contains the same elements as the set X.

If $x_1, x_2, \ldots, x_n$ are elements of X we write

$$X = \{x_1, x_2, \ldots, x_n, \ldots\}.$$

Most of the time this is used when we can extract the remaining elements of X from a pattern. For instance, while it's correct to write

$$\{x : x \text{ is the name of a US president } \} = \{\text{William}, \ldots\}$$

it would be a lot better to write

$$\{x : x \text{ is the name of a US president}\} = \{\text{William, George, Thomas, Dwight, Barrack, Grover}, \ldots\}$$

since it's easier to extract a pattern in the latter situation.

We think of $\{x_1, x_2, \ldots, x_n\}$ itself as a set based on the proposition

$$P(x) \text{ is true iff } x \text{ is equal some object listed.}$$

For instance,

$$\{\text{Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday}\} \tag{1}$$

is the set which corresponds to the proposition which is true iff the given object is the English name for a day of the week.

DEFINITION 2.1. We say that a set A is a *subset* of a set X (denoted $A \subseteq X$) and X is a *superset* of A (denoted $X \supseteq A$)

$$\forall a \in A \Rightarrow a \in X.$$

(In words every element of A is an element of X.) We also say that A is *contained* in X and X contains A.

In terms of propositions, this means that if $A(x)$ is true, then $X(x)$ is also true, i.e. the proposition corresponding to X is true more often than the proposition corresponding to A. For instance, the set listed in (1) is a subset of each of the sets

$$\{x : x \text{ is a word in the English langauge ending in the letter "y"}\},$$
$$\{x : x \text{ is a word showing up in this document}\},$$
$$\{x : x \text{ is an English word containing more than 4 letters}\}$$

tail since we are now using the notion of equality!

[II] We may also say that x is a *point* in X.

since, for instance, every day of the week is an English word ending in the letter y.

Definition 2.2. If we have that $A \subseteq B$ and $B \subseteq A$ for sets A and B, then we say that A and B are *equal* or are *the same sets*. As you would guess, if sets A and B are equal, we denote this by $A = B$. If $A \neq B$ if A and B are not equal.

While at first this definition may seem silly, it gives us an outline of how we show that two sets are actually equal in practice. Notice that the statement $A \neq B$ means

$$\exists a \in A \text{ s.t. } a \notin B \quad \text{or} \quad \exists b \in B \text{ s.t. } b \notin A.$$

Notation 2.3. The notation $A \subset X$ is shorthand for the statement that $A \subseteq X$ and $A \neq X$ (A not equal to X).

In terms of propositions, this means that the proposition (determining) A is true iff the proposition (determining) B is true. For instance, we have the equality

$$X = \{x : x \in X\}.$$

Exercise 2.1. (Russell's Paradox.) Let $\mathcal{A}$ be the collection of all sets which do not contain itself. That is

$$\mathcal{A} := \{A : A \text{ is a set and } A \notin A\}.$$

Assume that $\mathcal{A}$ is a set.

(1) Assume that $\mathcal{A} \in \mathcal{A}$. Prove that this implies that $\mathcal{A} \notin \mathcal{A}$.

(2) Now instead assume that $\mathcal{A} \notin \mathcal{A}$. Prove that this implies that $\mathcal{A} \in \mathcal{A}$.

(3) Conclude that you can't decide whether or not $\mathcal{A}$ is an element of itself and $\mathcal{A}$ can't possibly be a set.

Hint! Don't think too much about this. This is a "follow your nose" proof. It's an example of a proof that you simply cannot do by using your "right brain" and can cut through in 10 seconds by using your "left brain."

SOLUTION

Remark 2.4. This definition is both a flaw and a benefit of set theory because it means that the sets do not depend on the order that their elements are listed or if there are repeats. For instance, each of the sets

$$\{a, b, c\}, \qquad \{b, a, c\}, \quad \text{and} \quad \{a, a, c, b\}$$

are each equal (they each contain the same elements, a, b and c and no others).

Definition 2.5. The empty set is the set based on a proposition which is always false.

Such propositions exist, for instance, if we consider the set of all people in the Chicago Symphony Orchestra today who will be born after the year 20,000,000 AD, this set is empty (perhaps if you believe in reincarnation ...). Or we could be safer and take a proposition which is false by definition, for instance, we could take the set of people who are 18 years old and are older than 30. This set is denoted as $\varnothing$. By definition, the empty set is a set with no elements (it's the set consisting of objects which a proposition which is always false is true). A more mathematical way of defining the empty set is to define it as

$$\{x : x \neq x\}.$$

For instance, we have for all sets X, $\varnothing \subseteq X$. This is because if this inclusion did not hold, there would be an element of the empty set which is not in X. However the empty set contains no elements so it's false that $\varnothing$ is not a subset of X, so it is a subset of X.

Exercise 2.2. Show that $\varnothing$ does not depend on the false proposition used. (Like most proofs involving the empty set, your best bet is to prove this by contradiction). In other words if P and P' are always false, the sets that they determine (which you can use a temporary notation $\varnothing_P$ and $\varnothing_{P'}$) are equal.
Solution
Hint! Use the fact for any set X, $\varnothing \subseteq X$ as we just proved right before the start of this problem.

Definition 2.6. Suppose A and B are subsets of a set X. Then define:

(1) The *intersection* of A and B, $A \cap B$, as the set

$$\{x \in X : x \in A \text{ and } x \in B\}.$$

(2) The *union* of A and B, $A \cup B$, as the set

$$\{x \in X : x \in A \text{ or } x \in B\}.$$

(3) The *complement* of B in A, $A \setminus B$, as the set

$$\{x \in A : x \notin B\}.$$

Exercise 2.3. Prove
Solution

(1) $A, B \subseteq A \cup B$

(2) A and $B \subseteq Y \Leftrightarrow A \cup B \subseteq Y$

(3) $A, B \supseteq A \cap B$

(4) A and $B \supseteq Y \Leftrightarrow A \cap B \subseteq Y$.

THEOREM 2.7 (Basic Set Theory Laws). *Suppose that A, B and C are subsets of some set X. Then*

(1) Double Distributivity [III]

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

(2) DeMorgan

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

(3) Associativity

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

(4) Commutivity

$$A \cup B = B \cup A$$
$$A \cap B = B \cap A.$$

PROOF. These statements are just a translation of the usual logical statements. For instance, one can form a truth table which shows that for any three propositions, P, P_1, and P_2 that the propositions $P \vee (P_1 \wedge P_2)$ is equivalent (i.e. has the same truth table as) the proposition $(P \vee P_1) \wedge (P \vee P_2)$. Indeed, we could even save ourselves some work since if P is true than both are true, and if P is false then both are true iff both P_1 and P_2 are true. Translating this over to sets shows us that

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

as desired. The other parts are proved in the same way. ∎

With logic one should try to internalize such statements by coming up with everyday examples of how the statements apply. For instance, the statement[IV]

`Today I'm going outside and either going for a walk or a bike ride`

is clearly the same as the statement

`Today I'm going outside and going for a walk or today I'm going outside and taking a bike ride.`

[III] If you confuse the double distributive laws try replacing $\cup$ by multiplication and $\cap$ by addition in the first equation and the opposite in the second. Then in each case, the law looks just the usual distributive law.

[IV] Of course one need to wary of the distinction between how we often use the word in everyday language versus how a mathematician uses it. For instance, the statement I'm going shopping or I'm visiting my mother would be true if I did both! In math and in logic an or statement is false only if both parts are false. So the only way the statement I just made would be false is if I didn't go shopping and I didn't go to my mother's house!

DEFINITION 2.8.

(1) Let a and b be two objects. The *ordered pair* (a, b) is a new object which defined in such a way that $(a, b) = (a', b')$ holds iff $a = a'$ and $b = b'$.[V]

(2) More generally, if $a_1, \ldots, a_n$ are objects we define an *ordered n-tuple* of these objects as the new object

$$(a_1, \ldots, a_n)$$

which satisfies

$$(a_1, \ldots, a_n) = (a'_1, \ldots, a'_n) \Leftrightarrow a_1 = a'_1, \ldots, a_n = a'_n.$$

(3) Let A and B be sets. We form a new set, $A \times B$, which consists of all ordered pairs (a, b) for $a \in A$ and $b \in B$. This new set is called the *cartesian product* of A and B. More generally, if $A_1, \ldots, A_n$ are each sets we form a new set

$$A_1 \times \cdots \times A_n$$

which consists of all n-tuples $(a_1, \ldots, a_n)$ where $a_i \in A_i, i \in \overline{n}$. It is called the *cartesian product* of A_1 through A_n. This set can also be denoted as

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

If each of these sets, A_1 through A_n, are equal to some fixed set A we denote this cartesian product as A^n. Often, we will call elements of A^n *vectors* or *n-vectors*. This is especially the case when A has some notation of "addition." For instance, we shall often do this when A is the set of real numbers, $\mathbb{R}$. Occasionally we write an n–vector $(a_1, \ldots a_n)$ as a so called *column vector* [VI]

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

(4) The subset of A^n consisting of tuples of *distinct* points is denoted A^{*n}. Thus

$$A^{*n} = \{(a_1, \ldots, a_n) \in A^n : \text{ for } i, j \in \overline{n}, a_i = a_j \Leftrightarrow i = j\}.$$

For instance, as an ordered pair of natural numbers, $(2, 2) \in \mathbb{N}^2$ but $(2, 2) \notin \mathbb{N}^{*2}$ since $2 \neq 2$ is clearly false. For another example, $(7, 3, 4, 7) \in \mathbb{N}^4$, but, again, it is not an element of $\mathbb{N}^{*4}$ since the number 7 shows up twice.

REMARK 2.9. These definitions can be made more precise, see any foundations text book, for instance, {give a reference}. In Exercise 2.7 below I ask you to take the time to look up this definition. But again, for our purposes the property listed is what's more important. For instance, the ordered pairs are not equal

$$(\square, \star) \neq (\star, \square) \text{ even though the sets are } \{\square, \star\} = \{\star, \square\}.$$

In this way you can think of a set which contains two objects as being *unordered* pairs as opposed to the ordered pairs just defined.

[V] From a pedantic point of view, just because we defined something to have certain properties does not mean that such an object actually exists! See Remark 2.9 below.
[VI] One reason we do this is due to the fact that parenthesis are used to mean far too many things in math!

Remark 2.10. Technically, the three sets

$$(A \times A) \times A, \quad A \times (A \times A), \quad \text{and} \quad A \times A \times A$$

(the last being, by definition A^3) as not the same set (unless A is empty)! For an element $a \in A$ we have that $((a, a), a)$ is in the first set, (a, a, a) is in the second and $(a, (a, a))$ is in the third. However, we will only make a superficial distinction between them if we wish to emphasize something. We will more typically not make any distinction between them and as such we can be sloppy about these (trivial) differences. Making such a distinction is akin to making a distinction between two pairs of two things or four things.

Similarly, the two sets $A \times B$ and $B \times A$ are equal iff $A = B$ (c.f. Exercise 2.5). But again, we will not consider a situation where this distinction matters.[VII]

Warning! 2.11. Of course the cartesian product of sets has nothing to do with multiplication of numbers or the cross product of two vectors.

Exercise 2.4. Let A be any set. Prove that

$$\varnothing \times A = A \times \varnothing = \varnothing.$$

Solution

Hint! As with most problems involving the empty set you should prove this by contradiction.

Exercise 2.5. Let A, A', B, and B' be four non-empty sets (i.e. sets which are not the empty set and hence contain at least one element).

(1) Prove that

$$A \times B \subseteq A' \times B' \Leftrightarrow A \subseteq A' \wedge B \subseteq B'.$$

(2) Use the previous part to prove that

$$A \times B = A' \times B' \Leftrightarrow A = A', B = B'.$$

(3) Use the previous part to prove

$$A \times B = B \times A \Leftrightarrow A = B.$$

Solution

Hint! Let $A' = B$ and $B' = A$.

Exercise 2.6. The following statement (and hence its proof) is false. Find the error in the proof.[VIII]

Solution

[VII] Later, we will see in a clearer light what I mean by "matters."

FALSE THEOREM 2.12. *For sets A, A', and B, if $A \times B = A' \times B$, then $A = A'$.*

"**PROOF**".

$$\begin{aligned} A \times B = A' \times B &\Rightarrow A \times B/B = A' \times B/B \\ &\Rightarrow A \times \cancel{B/B} = A' \times \cancel{B/B} \\ &\Rightarrow A = A'. \end{aligned}$$

■

EXAMPLE 2.13.

(1) As you are no doubt aware, the set $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ is the usual (cartesian) plane and

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n\text{-times}}$$

is an n-dimensional space which in words is called "*are enn.*"

(2) If $A = \{1, 2, 3\}$, then

$$\begin{aligned} A^2 &= \{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2),(3,3)\}, \\ A^{*2} &= \{(1,2),(1,3),(2,1),(2,3),(3,1),(3,2)\}, \end{aligned}$$

and

$$A^{*3} = \{(1,2,3),(1,3,2),(2,1,3),(3,1,2),(2,3,1),(3,2,1)\}.$$

Of course, for any integer $n > 3$ we would have $A^{*n} = \varnothing$.

(3) It is useful to think of $A \times B$ as elements of a table. For instance, if $A = \{0, 1, 2\}$ and $B = \{\star, \square\}$, then $A \times B$ are the set of entries of the table

$\times$	0	1	2
$\star$	$(\star, 0)$	$(\star, 1)$	$(\star, 2)$
$\square$	$(\square, 0)$	$(\square, 1)$	$(\square, 2)$

.

EXERCISE 2.7. Look in your foundations book (math 215 at UIC) for a more technical definition of an ordered pair (a, b) and write it down here.

SOLUTION

[VIII] Your job in problems like this is to find the wrong step in the proof. Your job is **not** to explain why the theorem is false!

DEFINITION 2.14. Let X be a finite set. The *number of elements* in X is denoted $\#X$. If X is infinite, we write $\#X = \infty$.

RANT 2.15. (If you already know about functions, to come.) I'm sure you're dying to know the formal definition of a finite set and the number of elements in a set since we haven't formally defined how to count the number of elements in a set. Luckily, there are formal abstract definitions of these (I know, phew!). Namely one can define a set X to be finite if every injection $X \to X$ is a bijection.

One direction, which says that for X a finite set every injection is a bijection, is the Pigeon Hole Principle presented below in Theorem 2.55. The other direction, whose contrapositive says that for any infinite set X there exists an injection $X \to X$ which is not a bijection, comes by indexing a subset an infinite set X, with the positive integers $\mathbb{Z}^+$. This is simply the "everyday" way we think of an injection $\mathbb{Z}^+ \to X$. Such an injection can be shown to exists by successively "throwing away" one element of X at a time. This gives us a subset $X_{\mathbb{N}} := \{x_i\}_{i \in \mathbb{N}}$ of X. We then can then define an injection $X \to X$ via $x_i \mapsto x_{2i}$ for $x_i \in X_{\mathbb{N}}$ and $x \mapsto x$ for $x \notin X_{\mathbb{N}}$. This map can easily be checked to be injective, but it is not surjective since, for instance, x_1 is not in its image.

For more abstract purposes, this property of every injection being a bijection should be the "real" definition of the concept of finiteness we are taking for granted. Lastly, one can show that every finite set X, there exists a unique positive integer n so that X and $\overline{n}$ are in bijection. As such, we can formally define $\#X$ as this positive integer n.

$$\#X = n \Leftrightarrow \exists \text{ a bijection } X \to \overline{n}.$$

Even though we don't spend any more time on these formalities, which are best left for a foundations course, in practice we will nearly always end up showing that two sets has the same number of elements by proving there is a bijection between them. In this way, the number of elements a set has is a more primitive notion than counting the elements of the set. For instance, a farmer with 7 rocks will be know to plow 7 fields even if he/she cannot count to 7! They simply need to move one rock to a different pile every time they plow a field. This is discussed again in Exercise 2.29.

In our everyday experience, we have never seen anything that is not a set in some way.[IX] For instance, everything in life is the union of its fundamental particles. Or less abstractly, a state is the union of its people or its cities or its locations (depending on your perspective!). Thus it is natural to study sets themselves as objects. One way of doing this is with the power set.

DEFINITION 2.16. Given a set X, define the power set of X, $\mathcal{P}(X)$ as the set of all subsets of X. That is

$$\mathcal{P}(X) \stackrel{\text{def}}{=} \{A : A \subseteq X\}.$$

More generally, define $\mathcal{P}_n(X)$ as the set of all n-element subsets of X.

[IX] Physicists (well last time I checked!) postulate that electrons are not, but if you see an electron walking down the street, give me a call ... I have some students who like to write very small and I could use a hand grading their papers.

Example 2.17. If $X = \{1,2,3\} = \overline{3}$, then

$$\mathcal{P}_0(X) = \{\varnothing\},$$
$$\mathcal{P}_1(X) = \{\{1\},\{2\},\{3\}\},$$
$$\mathcal{P}_2(X) = \{\{1,2\},\{1,3\},\{2,3\}\},$$
$$\mathcal{P}_3(X) = \{\{1,2,3\}\} = \{X\},$$
$$\mathcal{P}(X) = \mathcal{P}_0(X) \cup \mathcal{P}_1(X) \cup \mathcal{P}_2(X) \cup \mathcal{P}_3(X) = \{\varnothing,\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\},X\}.$$

Of course, for any set X, $\mathcal{P}_0(X) = \{\varnothing\}$, and if $\#X = n$, then $\mathcal{P}_n(X)$ consists of the single object, the set X itself, i.e. $\mathcal{P}_n(X) = \{X\}$. Also note that we always have that the set $\mathcal{P}_1(X)$ is essentially the set X, the only difference is that every element of $\mathcal{P}_1(X)$ is wearing a fancy brace. We won't have the occasion to see the difference between these two sets so

Convention 2.18. Often we will be sloppy and consider $\mathcal{P}_1(X)$ (i.e. the subsets of X consisting of a single element) and X as the same. The identification being

$$\{x_1,\ldots,x_n,\ldots\} \leftrightsquigarrow \{\{x_1\},\ldots,\{x_n\},\ldots\}.$$

At the same time we will not make a distinction between a set containing a single element and the element itself. That is

$$x \leftrightsquigarrow \{x\}.$$

By definition, we have that $A \in \mathcal{P}(X)$ iff $A \subseteq X$. Also, the statement that $A_1,\ldots,A_n$ are each subsets of X (or are a family or collection of subsets of X) can be more susinctly written as saying that

$$\{A_1,\ldots,A_n\} \subseteq \mathcal{P}(X).$$

Notation 2.19. If $\#X = n$, then the number of elements in $\mathcal{P}_m(X)$ is denoted, by definition, as

$$\binom{n}{m}.$$

It is read *n choose m*.

Of course this number doesn't depend on what n things the set X consists! See the formula below to reaffirm this. It is the number of ways one can choose m things from X (disregarding order); since a choice of m things from X is the same thing as picking a subset of X which contains m elements, i.e. an element of $\mathcal{P}_m(X)$. For instance, in the set of vowels, {a,e,i,o,u}, the subset {a,e} is a certain choice of two vowels. It is well known that

$$\binom{n}{m} = \frac{n!}{m! \cdot (n-m)!}. \tag{2}$$

For instance, in Example 2.17, we verified this formula above in the special case that $n = 3$ and $m = 2$. Much later, in Example 10.96**(10.96)** on page 554 we will give a sweet group theory proof of this formula.

Exercise 2.8. Prove that

$$\binom{n}{2} = \sum_{i=1}^{n-1} i.$$

SOLUTION

Hint! It's a common anecdote that Gauss was able to do the following problem while he was still in elementary school. Check out

http://www.sigmaxi.org/amscionline/gauss-snippets.html

for over a 109 tellings of this (more than likely tall) tale.

Exercise 2.9. For sets $X_1, X_2 \subseteq X$ and $Y_1, Y_2 \subseteq Y$ prove that

SOLUTION

(1) $(X_1 \cup X_2) \times Y = X_1 \times Y \cup X_2 \times Y$.

(2) $X \times (Y_1 \cup Y_2) = X \times Y_1 \cup X \times Y_2$.

(3) $(X_1 \cap X_2) \times Y = X_1 \times Y \cap X_2 \times Y$.

(4) $X \times (Y_1 \cap Y_2) = X \times Y_1 \cap X \times Y_2$.

(5) $\mathcal{P}(X_1) \subseteq \mathcal{P}(X)$.

(6) $\mathcal{P}(X_1) \cap \mathcal{P}(X_2) = \mathcal{P}(X_1 \cap X_2)$.

(7) Give an example which shows that in general,

$$\mathcal{P}(X_1 \cup X_2) \neq \mathcal{P}(X_1) \cup \mathcal{P}(X_2).$$

Exercise 2.10. Suppose that X and Y are finite sets. Prove that

SOLUTION

$$\#\mathcal{P}(X) = 2^{\#X} \text{ and } \#(X \times Y) = \#X \cdot \#Y$$

Hint! Probably the best way to show this is by induction. In the second equality you should deal with the special case that $\#Y = 1$ first.

Exercise 2.11. **T** or **F**, $\mathcal{P}_2(\mathbb{N}) \subseteq \mathcal{P}_3(\mathbb{N})$.

SOLUTION

DEFINITION 2.20.

(1) A *collection* of subsets of a set X is just a subset of the power set of X, $\mathcal{P}X$.

(2) Let $\mathcal{A}$ be a collection of subsets of a set X. Then define:

(a) The intersection over $\mathcal{A}$

$$\bigcap_{A\in\mathcal{A}} A \stackrel{\text{def}}{=} \{a : a \in A \text{ for all } A \in \mathcal{A}\}.$$

(b) The union over $\mathcal{A}$

$$\bigcup_{A\in\mathcal{A}} A \stackrel{\text{def}}{=} \{a : a \in A \text{ for some } A \in \mathcal{A}\}.$$

With these definitions, one could prove a theorem similar to Theorem 2.7. However, if you glance at your foundations text where they no doubt prove such a thing, you can't just use this theorem. This is because the unions and intersections over infinite collections behave quite differently then their corresponding version for infinite collections.

EXAMPLE 2.21. Let $\mathcal{A}$ be the collection of subsets of $\mathbb{R}$ of the form

$$\left(-\frac{1}{n}, \frac{1}{n}\right)$$

for some positive integer n. Then

$$\bigcap_{A\in\mathcal{A}} A = \{0\}.$$

This is because for any positive integer n, 0 is between $-\frac{1}{n}$ and $\frac{1}{n}$, but for any other real number r there exists an integer n so that $1/n$ is smaller, in absolute value, than r is.

PROPOSITION 2.22.

(1) *If $\mathcal{A}$ is a non-empty collection of subsets of X each containing a set Y, then*

$$Y \subseteq \bigcap_{A\in\mathcal{A}} A$$

as well.

(2) *If some $A \in \mathcal{A}$ is contained in a set B for some $i \in \mathcal{A}$, then*

$$\bigcap_{A\in\mathcal{A}} A \subseteq B$$

as well. (*This is used in the special case that $B = A$ for some $A \in \mathcal{A}$.*)

Proof. Let $a \in Y$ be arbitrary. Then since $Y \subseteq A$ for all $A \in \mathcal{A}$ we have that $a \in A$ for all $A \in \mathcal{A}$. Hence

$$a \in \bigcap_{A \in \mathcal{A}} A.$$

Since a was arbitrary, we have

$$Y \subseteq \bigcap_{A \in \mathcal{A}} A$$

as desired. The proof of the second part is left as an exercise. ∎

Exercise 2.12. Let X be a non-empty set and x an element of X. Let $\mathcal{A}$ be the collection of all subsets of X which contain a given element x.[X] Prove that

(1) $\{x\} = \bigcap_{A \in \mathcal{A}} A$;

(2) $X = \bigcup_{A \in \mathcal{A}} A$.

[X] Symbolically,

$$\mathcal{A} := \{A \subseteq X : x \in A\}.$$

2.2 Relations in General

One cannot possibly study mathematics without studying functions and equivalence relations (generalized equality).

Relations are an incredibly primitive concept which will be used an intermediately between either functions or equivalence relations. Typically abstract relations have only a few purposes:

- They allow us to define many things at once, e.g. functions, posets, and equivalence relations.
- They have so little structure that proofs are typically 100% left brained ... i.e. proofs that stress proof structure more than content in the sense that every step leads naturally to the next step if one has a solid working grasp on the basic definitions involved. In this way they are what I think of as "left brain exercise" which is often the greatest obstacle for a new, budding mathematician.
- Functions are inherently asymmetric. This is because if f is a function, then f^{-1} most often is **not** a function. However, f^{-1} is always a relation.

I in now way recommend that a student get lost in this section. Again, like much of the first, it should be though of as a curiosity and a bit of exercise. If you do not find the section interesting, then like the athlete who hates running, I recommend that you go for a bike ride, I mean skip it.

Definition 2.23. Let X and Y be sets. A subset of a $X \times Y$ is called a *relation from a set X to Y*. A relation from X to X is a relation *on* X.

Thus a relation has three bits of information to it. A set X, a set Y, and a subset $\mathcal{R} \subseteq X \times Y$. We summarize all of this by writing $\mathcal{R} \colon X \rightsquigarrow Y$. We call X the *first set* of the relation and Y the *second set* of the relation. Often, we refer to the actual subset $\mathcal{R}$ itself as the relation. However, strictly speaking, simply writing the subset omits two bits of information, the set X and the set Y.

Definition 2.24. Suppose $\mathcal{R} \subseteq X \times Y$ is a relation. If $(a, b) \in \mathcal{R}$ we say that a *is $\mathcal{R}$–related* or simply *related* to b. We often denote this as $a \mathcal{R} b$ instead of $(a, b) \in \mathcal{R}$. Occasionally, we also write this as $a \overset{\mathcal{R}}{\mapsto} b$.

It should be noted there there is no creativity involved in defining relations. Simply write out every element of $X \times Y$ and then circle your favorite elements to define your very own!

SubDefinition 2.25.

(1) For any set X,

$$\mathbf{1}_X \overset{\text{def}}{=} \{(a, a) : a \in X\}$$

is a relation on X called the *identity*.

(2) If $A \subseteq X$, we get a relation from A to X, $\iota_A^X \colon A \rightsquigarrow X$ (or just ι_A, or even ι) defined by

$$\mathbf{1}_X \cap (A \times X).$$

This is called the *inclusion* of A into X.

Thus we have

$$(a, b) \in \mathbf{1}_X \Leftrightarrow a = b.$$

Also note that as sets, the inclusion ι_A^X is the same set $\mathbf{1}_A = \{(a, a) : a \in A\}$. However the second set of the inclusion relation is X, not A as it is for the identity.

	x_1	x_2	x_3	x_4	x_5
x_1	(x_1, x_1)	(x_1, x_2)	(x_1, x_3)	(x_1, x_4)	(x_1, x_5)
x_2	(x_2, x_1)	(x_2, x_2)	(x_2, x_3)	(x_2, x_4)	(x_2, x_5)
x_3	(x_3, x_1)	(x_3, x_2)	(x_3, x_3)	(x_3, x_4)	(x_3, x_5)
x_4	(x_4, x_1)	(x_4, x_2)	(x_4, x_3)	(x_4, x_4)	(x_4, x_5)
x_5	(x_5, x_1)	(x_5, x_2)	(x_5, x_3)	(x_5, x_4)	(x_5, x_5)

	x_1	x_2	x_3	x_4	x_5
x_1	(x_1, x_1)	(x_1, x_2)	(x_1, x_3)	(x_1, x_4)	(x_1, x_5)
x_2	(x_2, x_1)	(x_2, x_2)	(x_2, x_3)	(x_2, x_4)	(x_2, x_5)
x_3	(x_3, x_1)	(x_3, x_2)	(x_3, x_3)	(x_3, x_4)	(x_3, x_5)
x_4	(x_4, x_1)	(x_4, x_2)	(x_4, x_3)	(x_4, x_4)	(x_4, x_5)
x_5	(x_5, x_1)	(x_5, x_2)	(x_5, x_3)	(x_5, x_4)	(x_5, x_5)

Figure 2.1: The identity on the set $X := \{x_1, x_2, x_3, x_4, x_5\}$ highlighted in yellow. If $A := \{x_1, x_2\}$, then the inclusion, ι_A^X, is highlighted in green.

EXAMPLE 2.26. Let

$$X = \{\clubsuit, \spadesuit, \heartsuit, \diamondsuit\} \qquad Y = \{\bigstar, \blacklozenge, \blacktriangleright\}.$$

Then there are $4 \cdot 3 = 12$ elements in the set $X \times Y$, and so there are a total of $2^{\#(X \times Y)} = 2^{12} = 4,096$ different relations defined from X to Y. Here are 3:

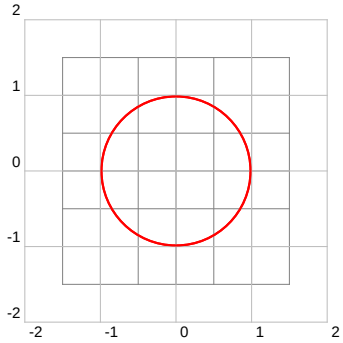
$$\mathcal{R}_1 := \{(\clubsuit, \blacklozenge), (\spadesuit, \bigstar)\}$$
$$\mathcal{R}_2 := \{(\heartsuit, \bigstar), (\spadesuit, \bigstar), (\diamondsuit, \blacktriangleright), (\heartsuit, \blacktriangleright)\}$$
$$\mathcal{R}_3 := \{(\diamondsuit, \blacktriangleright), (\diamondsuit, \bigstar), (\diamondsuit, \blacklozenge)\}.$$

In the second relation, $\heartsuit$ is related to $\bigstar$, but not in the first and third relations.

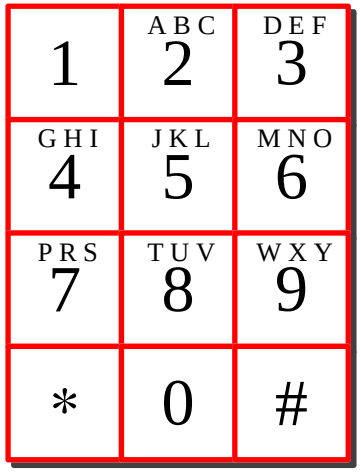
EXAMPLE 2.27. The set

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$$

defines a relation on the real numbers $\mathbb{R}$. As a subset of the plane, it is the usual unit circle, $\mathbb{S}^1$. In this relation, for example, we have that 1 is related only to 0, 1/2 is related to both $\frac{\sqrt{3}}{2}$ and $-\frac{\sqrt{3}}{2}$, etc. In fact everything anybody has ever written on a piece of paper defines a relation on $\mathbb{R}$ (after choosing a coordinate system on your paper.) Simply relate pairs of points that are covered with lead (or ink, or crayon, etc).



Example 2.28. A more familiar relation could be obtained by defining X as the set of integers 0 through 9 and defining Y to be the set of letters in the English alphabet. Then we define a relation $\mathcal{R}$ by stating that a number if related to a letter iff that letter is written on the given number on a standard telephone's keypad. For instance, the number 1 is not related to any letter, the number 2 is related to the letters A,B, and C. No number is related to Q or Z, etc.



Example 2.29. If we use the calc I definition of a function, for a function f with domain X, the graph of f

$$\{(x, f(x)) : x \in X\}$$

defines a relation. Indeed, every advanced math textbook agrees that the usual, elementary definition of a function sucks. We will therefore give our non-sucky definition of a function as a relation with certain properties. More generally, every equation determines a relation. For instance, the relation $x^2 + y^2 = 1$ determines the relation

$$\Gamma := \{(x, y) : x^2 + y^2 = 1\} \subseteq \mathbb{R}^2$$

we first discussed in Example 2.27. Of course this relation is not the graph of any function. A pre-calculus student would say that it fails the horizontal line test since $(0, 1)$ and $(0, -1)$ are both elements of Γ.

Exercise 2.13. Prove that there are a total of $2^{\#X \cdot \#Y}$ relations from X to Y. How many relations are there from Y to X? Why?

Solution

A useful tool (when X and Y are finite) for relations is to form the *digraph* associated with the relation. This is drawn as follows:

- On one side pot a dot for every element of X and on the other side dots corresponding to elements of Y.
- Connect the dot corresponding to the element $x \in X$ to the element of $y \in Y$ by an arrow (pointing from x's dot to y's) iff $x \mathcal{R} y$.

Example 2.30. Figure 2.2 depicts a relation where a is related to 2 as well as 3. Also, in this relation nobody is related to to 4, and d is related to nobody.

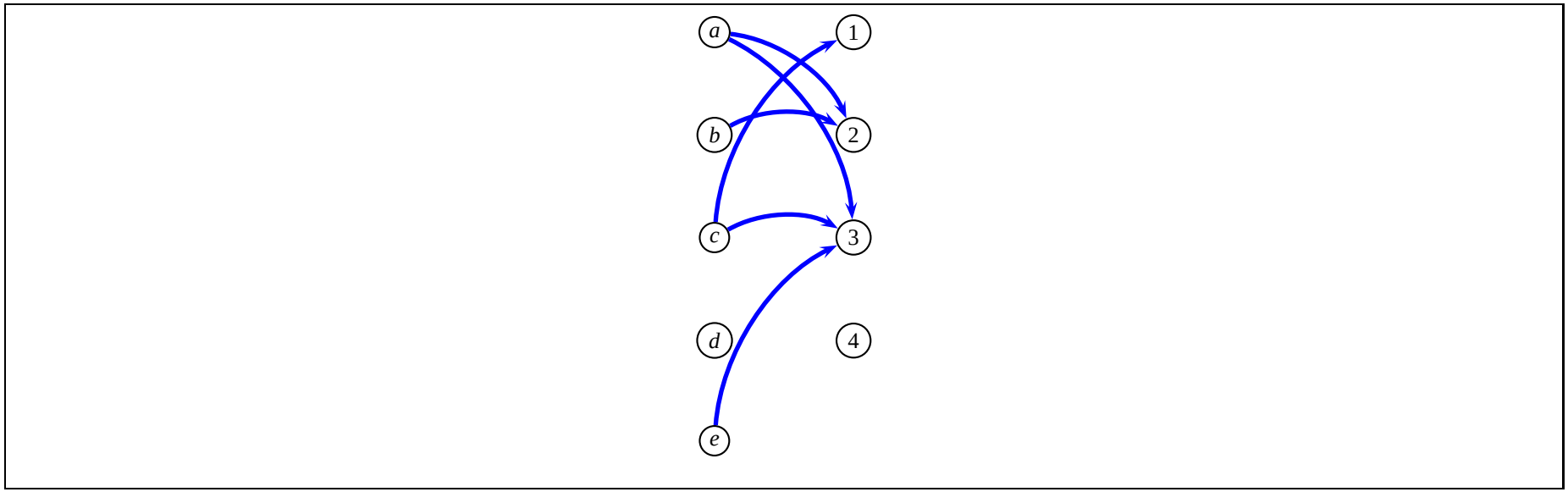


Figure 2.2: Figure for Example 2.30

EXAMPLE 2.31. For relations on X we do this by drawing elements of X only once (therefore loops are possible). For instance, Figure 2.3 depicts the relation on the set $\overline{6} = \{1, 2, 3, 4, 5, 6\}$ given by

$$1\,\mathcal{R}\,1 \qquad 6\,\mathcal{R}\,3 \qquad 4\,\mathcal{R}\,4 \qquad 3\,\mathcal{R}\,3 \qquad 5\,\mathcal{R}\,1$$
$$1\,\mathcal{R}\,4 \qquad 1\,\mathcal{R}\,5 \qquad 2\,\mathcal{R}\,6 \qquad 4\,\mathcal{R}\,1$$

or, if you prefer it written as a subset, it is

$$\mathcal{R} := \{(1,1),(6,3),(4,4),(3,3),(5,1),(1,4),(1,5),(2,6),(4,1)\} \subseteq \overline{6} \times \overline{6}.$$

EXERCISE 2.14.

(1) Define a relation $\overline{4} \rightsquigarrow \overline{4}$ by saying that two elements $a, b \in \overline{4}$ are related iff $|a - b| = 2$ or 0. Explicitly write down all elements of this relation.

(2) Same problem as above but just replace $\overline{4}$ by $\overline{5}$.

SOLUTION

DEFINITION 2.32. Let $\mathcal{R}\colon X \rightsquigarrow Y$ be a relation.

(1) Let $\mathcal{S}$ be another relation. We say that $\mathcal{S}$ and $\mathcal{R}$ are *composable* if the first set of $\mathcal{S}$ is the second set of $\mathcal{R}$, Y.

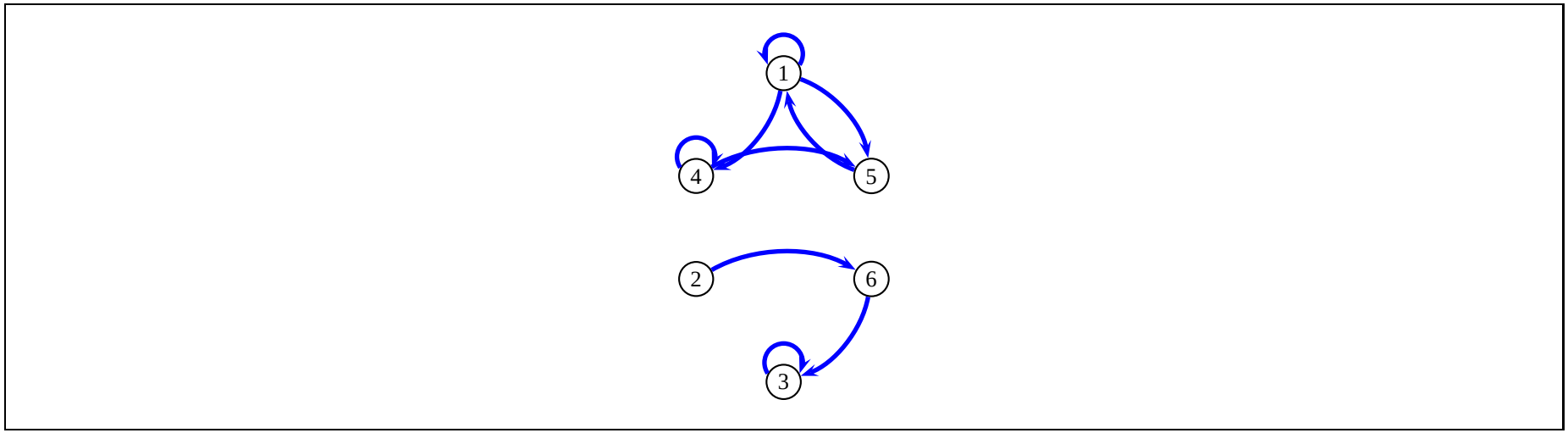


Figure 2.3: Figure for Example 2.31

(2) Suppose that $\mathcal{S}$ and $\mathcal{R}$ are composable. We define $\mathcal{S} \circ \mathcal{R}$ as the relation

$$\{(x, z) : \exists y \in Y \text{ s.t. } (x, y) \in \mathcal{R} \text{ and } (y, z) \in \mathcal{S}\}.$$

This is read as *the composition* (or the composite) of $\mathcal{S}$ and $\mathcal{R}$.

(3) The relation $\mathcal{R}^{-1}$ is the relation $Y \rightsquigarrow X$ given by $\{(x, y) : (y, x) \in \mathcal{R}\}$.

REMARK 2.33. Perhaps the composition of two relations looks nicer without the existential quantifier:

$$x \mathcal{R} y \quad \text{and} \quad y \mathcal{S} z \Rightarrow x \mathcal{S} \circ \mathcal{R} z,$$

or said in the membership notation

$$(x, y) \in \mathcal{R} \quad \text{and} \quad (y, z) \in \mathcal{S} \Rightarrow (x, z) \in \mathcal{S} \circ \mathcal{R}.$$

In this way you should think of y as an intermediate which allows x to be related to z.

Unlike functions, every relation is invertible. In terms of the digraph, we are simply turning around all the arrows. But like functions, if we are given two relations, and if the second set of the first relation does not match the first set of the second the two relations **cannot** be composed. The notion doesn't even make sense.[1]

[1] It would be like taking a flight from Madrid to Paris and then taking a connecting flight from Sidney to New York City. It's just gibberish.

EXERCISE 2.15. Let

$$X = \{\clubsuit, \spadesuit, \heartsuit, \diamondsuit\} \quad \text{and} \quad Y = \{\bigstar, \blacklozenge, \blacktriangleright\}.$$

Let $\mathcal{R}$ be the relation depicted in Figure 2.4.

SOLUTION

(1) Write $\mathcal{R}$ as a subset of $X \times Y$.

(2) Calculate $\mathcal{R}^{-1}$.

(3) Calculate $\mathcal{R} \circ \mathcal{R}^{-1}$ and $\mathcal{R}^{-1} \circ \mathcal{R}$. Are either equal to (the relevant) identity?

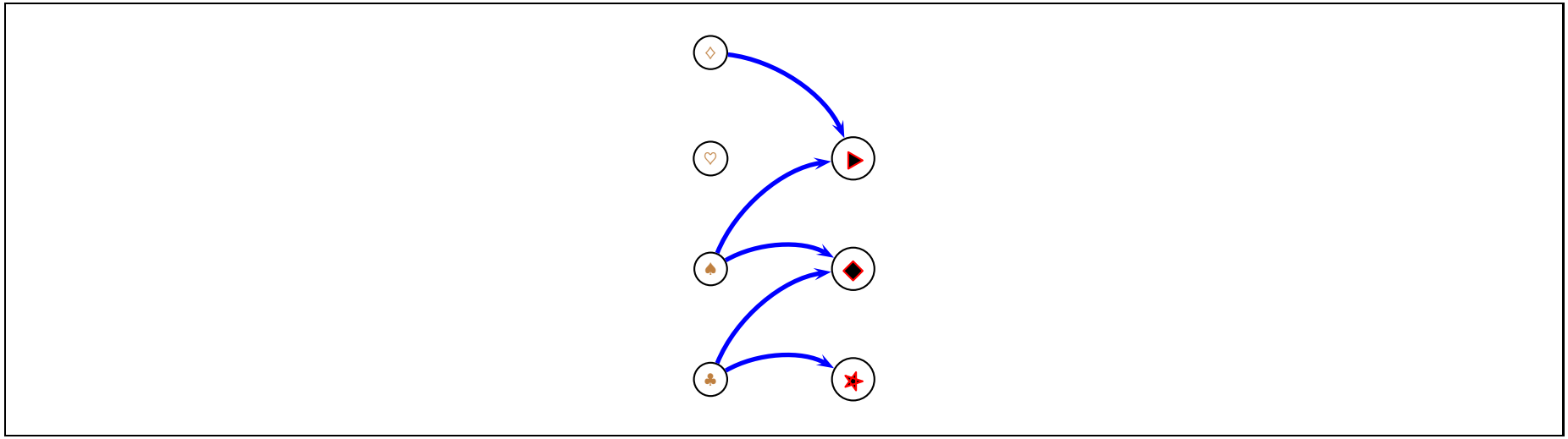


Figure 2.4: Figure for Exercise 2.15

EXERCISE 2.16. Let **P** denote the set of all people.

SOLUTION

(1) Define a relation $\mathcal{R}$ on **P** which relates person 1 to person two iff person 1 is a parent of person 2, i.e. $p_1 \mathcal{R} p_2$ iff p_2 is a parent of p_1.

(a) Describe each of the following relations in words:

(i) $\mathcal{R} \circ \mathcal{R}$

(ii) $\mathcal{R}^{-1}$

(iii) $\mathcal{R} \circ \mathcal{R}^{-1}$

(iv) $\mathcal{R}^{-1} \circ \mathcal{R}$.

(b) Is $\mathcal{R}$ defined everywhere?

(c) Is $\mathcal{R}$ surjective?

(2) Now suppose that we define another relation, $\mathcal{S}$, on **P** which relates person 1 to person 2 iff person 1 is the brother of person 2. Describe the following in words:

(a) $\mathcal{R} \circ \mathcal{S}$

(b) $\mathcal{S} \circ \mathcal{R}$

(c) $\mathcal{S} \circ \mathcal{R}^{-1}$.

Exercise 2.17. For any relation $\mathcal{R}\colon X \rightsquigarrow Y$ prove

$$\mathcal{R} \circ \mathbf{1}_X = \mathcal{R} = \mathbf{1}_Y \circ \mathcal{R}.$$

Solution

Hint! You are trying to prove the equality of sets.

Exercise 2.18. Suppose that $\mathcal{R}$, $\mathcal{S}$, and $\mathcal{T}$ are composable relations (i.e. the first set of $\mathcal{S}$ is the second set of $\mathcal{R}$ and the second set of $\mathcal{S}$ is the first set of $\mathcal{T}$). Show that both

$$\mathcal{T} \circ (\mathcal{S} \circ \mathcal{R}) \quad \text{and} \quad (\mathcal{T} \circ \mathcal{S}) \circ \mathcal{R}$$

are equal to (and in particular they are equal to each other!)

$$\{(a, d) : \exists b, c \text{ s.t. } (a, b) \in \mathcal{R}, (b, c) \in \mathcal{S}, (c, d) \in \mathcal{T}\}.$$

Generalize this to show that in a sequence of n-relations one never needs parenthesis in any composition. This is known as *generalized associativity*.

Solution

Exercise 2.19. Let $\mathcal{R}_i\colon X \rightsquigarrow Y$, $\mathcal{S}_i\colon Y \rightsquigarrow Z$, $(i = 1, 2)$ be relations. Prove that

(1) $(\mathcal{S}_1 \cup \mathcal{S}_2) \circ (\mathcal{R}_1 \cup \mathcal{R}_2) = \mathcal{R}_1 \circ \mathcal{S}_1 \cup \mathcal{R}_1 \circ \mathcal{S}_2 \cup \mathcal{R}_2 \circ \mathcal{S}_1 \cup \mathcal{R}_2 \circ \mathcal{S}_2$;

(2) Show that only 1 inclusion of **(1)** holds if we swap $\cup$ with $\cap$.

(3) $(\mathcal{R} \cup \mathcal{S})^{-1} = \mathcal{R}^{-1} \cup \mathcal{S}^{-1}$.

(4) $(\mathcal{R} \cap \mathcal{S})^{-1} = \mathcal{R}^{-1} \cap \mathcal{S}^{-1}$.

Solution

Note that the first set of $\mathcal{S} \circ \mathcal{R}$ is the first set of $\mathcal{R}$ and the second set of this composition is the second set of $\mathcal{S}$. In terms of the diagram we follow two edges. For instance, in the digraph above we have that

$$\mathcal{R} \circ \mathcal{R} = \{(1, 1), (1, 4), (1, 5), (2, 3), (3, 3), (4, 1), (4, 4), (4, 5), (5, 1), (5, 4), (5, 5), (6, 3)\}.$$

For any relation $\mathcal{R}$ on a set X with n elements, $\{x_1, \ldots, x_n\}$ (most of the time this will be the set $\overline{n}$) define an n by n matrix, called the adjacency matrix of $\mathcal{R}$, denoted $\widehat{\mathcal{R}}$, by declaring that

$$\hat{\mathcal{R}}_{ij} = 1 \Leftrightarrow x_i \mathcal{R} x_j.$$

EXERCISE 2.20.

(1) Prove that

$$(\widehat{\mathcal{R}})^\top = \widehat{\mathcal{R}^{-1}}$$

(for the definition of the transpose of a matrix jump ahead to see Definition 3.209 coming up on page 193).

(2) For a matrix M with integer entries define $M^\bigstar$ to be the matrix which replaces every non-zero entry with a 1. Prove that

$$\left(\widehat{\mathcal{R}} \cdot \widehat{\mathcal{S}}\right)^\bigstar = \widehat{\mathcal{R} \circ \mathcal{S}}.$$[II]

SOLUTION

An instance of the definition from the last exercise, if we take $\mathcal{R}$ as in the example from the figure we have that

$$\widehat{\mathcal{R}} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

When we square this matrix we get

$$\widehat{\mathcal{R}^2} = \begin{pmatrix} 3 & 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 2 & 2 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

For instance, the 3 in the $(4, 1)$ entry corresponds to the three paths from 4 to 3 which take 2 steps to complete. Indeed, you could

(1) Stay at 4 (possible since 4 is related to itself) then go to 1.

(2) Go from 4 to 5, then stay at 1.

(3) Go from 4 to 5, then go from 5 to 1.

[II] Actually the entry $\left(\widehat{\mathcal{R}} \cdot \widehat{\mathcal{S}}\right)_{i,j}$ denotes the number of k's with $(x_i, x_k) \in \mathcal{S}$ and $(x_k, x_j) \in \mathcal{R}$, which, by definition, must be at least one if x_i is to be related to x_j in the composition.

Definition 2.34. Suppose that $\mathcal{R}\colon X \rightsquigarrow Y$ is a relation, $A \subseteq X$ and $B \subseteq Y$.

(1) The *image* of A, denoted $\mathcal{R}(A)$, under $\mathcal{R}$ as the set

$$\{y \in Y : \exists a \in A \text{ s.t. } (a,y) \in \mathcal{R}\}.$$

(2) The *inverse image* of B under $\mathcal{R}$ is defined to be the image of B under $\mathcal{R}^{-1}$.[III] In symbols it is the set

$$\{x \in X : \exists b \in B \text{ s.t. } (x,b) \in \mathcal{R}\}.$$

Notice the notation for this is $\mathcal{R}^{-1}(B)$[IV].

(3) The *image of* $\mathcal{R}$ (or the *range* of $\mathcal{R}$), denoted as $\operatorname{Im}(\mathcal{R})$ is the image of the entire set X under $\mathcal{R}$. That is

$$\operatorname{Im}(\mathcal{R}) \stackrel{\text{def}}{=} \mathcal{R}(X).$$

In more detail, the image is the set

$$\{y \in Y : \exists x \in X \text{ s.t. } (x,y) \in \mathcal{R}\}.$$

(4) The *domain of* $\mathcal{R}$, denoted $\operatorname{Dom}(\mathcal{R})$, is the inverse image of Y under $\mathcal{R}$. That is

$$\operatorname{Dom}(\mathcal{R}) \stackrel{\text{def}}{=} \mathcal{R}^{-1}(Y).$$

In more detail, the domain is the set

$$\{x \in X : \exists y \in Y \text{ s.t. } (x,y) \in \mathcal{R}\}.$$

In terms of the digraph the domain is simply the set of elements of the first set which have some arrow coming from it and the image of is the set of guys who have some arrow going to it. The image of a subset is the

Exercise 2.21. For any two relations $\mathcal{R}$ and $\mathcal{S}$ whose composition is defined, prove that for any subset, A, of the first set of $\mathcal{R}$ we have that

$$(\mathcal{S} \circ \mathcal{R})(A) = \mathcal{S}(\mathcal{R}(A)).$$

Solution

Exercise 2.22. Let X be a set and $A \subseteq X$. Prove for any $B \subseteq X$ that $\left(\iota_A^X\right)^{-1}(B) = A \cap B$, where ι_A^X is the inclusion of A into X.

Solution

Definition 2.35. Let $\mathcal{R}\colon X \rightsquigarrow Y$ be a relation. Then

(1) $\mathcal{R}$ is *defined everywhere* if for any $x \in X$ there exists a (at least one) $y \in Y$ so that $(x, y) \in \mathcal{R}$. That is if $\text{Dom}(\mathcal{R}) = X$ or, said another way, every element of X is related to at least one element of Y.

(2) $\mathcal{R}$ is *surjective* (or *onto*) if for any $y \in Y$ there exists an $x \in X$ so that $(x, y) \in \mathcal{R}$. That is if $\text{Im}(\mathcal{R}) = Y$.

DEFINITION 2.36.

(1) $\mathcal{R}$ is *single-valued* if for any $x \in X$ we have that

$$x \mathcal{R} y_1 \text{ and } x \mathcal{R} y_2 \Rightarrow y_1 = y_2.$$

Said in another way, this means that every element of X is related to at most one element of Y.

(2) $\mathcal{R}$ is *injective* (or *one-to-one*) if for any $y \in X$ we have that

$$x_1 \mathcal{R} y \text{ and } x_2 \mathcal{R} y \Rightarrow x_1 = x_2.$$

EXERCISE 2.23. Let $\mathbf{M}$ be a matrix. Define $R(\mathbf{M})_i$ (resp. $C(M)_i$) as the sum of entries in the i^{th} row of $\mathbf{M}$ (resp. column of $\mathbf{M}$). Now let $\mathcal{R}$ be a relation whose first and second sets are the same and $\widehat{\mathcal{R}}$ denote its adjacency matrix from Exercise 2.20. Show that for all positive integers i less than or equal to the number of elements of the first set of $\mathcal{R}$ that we have

(1) $C(\mathcal{R})_i \leq 1$ iff $\mathcal{R}$ is injective,

(2) $C(\mathcal{R})_i \geq 1$ iff $\mathcal{R}$ is surjective,

(3) $R(\mathcal{R})_i \geq 1$ iff $\mathcal{R}$ is defined everywhere,

(4) $R(\mathcal{R})_i \leq 1$ iff $\mathcal{R}$ is single-valued.

SOLUTION

Despite what you would expect from functions, we have seen that for most relations, $\mathcal{R}$, we do **not** have that $\mathcal{R} \circ \mathcal{R}^{-1} = \mathbf{1}$.[V] The following derives a condition when this does happen.

EXAMPLE /THEOREM 2.37. For any surjective, single-valued relation $\mathcal{R}$ we have $\mathcal{R} \circ \mathcal{R}^{-1} = \mathbf{1}$.

PROOF. We are trying to show the equality of the two sets, $\mathcal{R} \circ \mathcal{R}^{-1}$ and $\mathbf{1}$.

$\supseteq$ Let $(a, a) \in \mathbf{1}$. Since $\mathcal{R}$ is surjective there exists an x in the first set of $\mathcal{R}$ so that $(x, a) \in \mathcal{R}$. But, by definition, this means that $(a, x) \in \mathcal{R}^{-1}$. Sticking these together shows us that $\mathcal{R} \circ \mathcal{R}^{-1}$. Since (a, a) is an arbitrary element of $\mathbf{1}$ we have $\mathbf{1} \subseteq \mathcal{R} \circ \mathcal{R}^{-1}$ like we wanted.

[V] In fact this happens iff both $\mathcal{R}$ and its inverse $\mathcal{R}^{-1}$ are functions.

$\subseteq$ Now let $(a,b) \in \mathcal{R} \circ \mathcal{R}^{-1}$. In order to show that $(a,b) \in \mathbf{1}$ we need to show that $a = b$. By definition, if

$$\begin{aligned}(a,b) \in \mathcal{R} \circ \mathcal{R}^{-1} &\Rightarrow \exists c \text{ s.t. } (a,c) \in \mathcal{R}^{-1}, (c,b) \in \mathcal{R} && \text{(Def of comp)}\\ &\Rightarrow \exists c \text{ s.t. } (c,a) \in \mathcal{R}, (c,b) \in \mathcal{R} && \text{(Def of inv of a relation)}\\ &\Rightarrow a = b. && (\mathcal{R} \text{ is single-valued})\end{aligned}$$

But this is exactly what we wanted to prove: since $a = b$, $(a,b) \in \mathbf{1}$ proving the reverse inclusion is also valid. ■

EXERCISE 2.24. For any two relations $\mathcal{R}$ and $\mathcal{S}$ show that $\mathcal{R} \subseteq \mathcal{S}$ iff $\mathcal{R}^{-1} \subseteq \mathcal{S}^{-1}$.

SOLUTION

EXERCISE 2.25. Suppose that $\mathcal{R}, \mathcal{S} \colon X \rightsquigarrow Y$. Prove that

$$\mathcal{R} \subseteq \mathcal{S} \Leftrightarrow \forall A \subseteq X,\ \mathcal{R}(A) \subseteq \mathcal{S}(A).$$

SOLUTION

Careful! You probably can't do both directions at the same time. The forward direction is 100% left-brained, the reverse takes a bit of trickery since it may seem that you will not need the full strength of the assumption.

EXERCISE 2.26. For any relations, $\mathcal{R} \colon X \rightsquigarrow Y$ and $\mathcal{S} \colon Y \rightsquigarrow Z$, prove that

(1) If $\mathcal{R}$ and $\mathcal{S}$ are single valued, then so is $\mathcal{S} \circ \mathcal{R}$,

(2) If $\mathcal{R}$ and $\mathcal{S}$ are defined everywhere, then so is $\mathcal{S} \circ \mathcal{R}$,

(3) If $\mathcal{R}$ and $\mathcal{S}$ are surjective, then so is $\mathcal{S} \circ \mathcal{R}$,

(4) If $\mathcal{R}$ and $\mathcal{S}$ is injective, then so is $\mathcal{S} \circ \mathcal{R}$.

SOLUTION

EXERCISE 2.27. For any relation, $\mathcal{R} \colon X \rightsquigarrow Y$, show that

$$\begin{aligned}\mathcal{R} \circ \mathcal{R}^{-1} \subseteq \mathbf{1}_X &\Leftrightarrow \mathcal{R} \text{ is single-valued}\\ \mathcal{R} \circ \mathcal{R}^{-1} \supseteq \mathbf{1}_X &\Leftrightarrow \mathcal{R} \text{ is onto}\\ \mathcal{R}^{-1} \circ \mathcal{R} \subseteq \mathbf{1}_X &\Leftrightarrow \mathcal{R} \text{ is one to one}\\ \mathcal{R}^{-1} \circ \mathcal{R} \supseteq \mathbf{1}_X &\Leftrightarrow \mathcal{R} \text{ is defined everywhere.}\end{aligned}$$

Hint! You only need to prove two of these statements if you understand the definitions well.

SOLUTION

EXAMPLE 2.38. Let $\mathcal{R}$ be the relation defined by the digraph in Figure 6. Looking at the digraph makes it easy to see the properties of a relation. For instance, to check whether or not $\mathcal{R}$ is single-valued we simply need to check that no point in the first set $\mathcal{R}$ has more than one arrow coming from it. We see this is the case for $2, 3, 5$, and 6 but it is not the case at 1 and 4 (in particular, that relation is not single-valued). To see if the relation is defined everywhere, we need at least one arrow leaving from each vertex for onto we need at least one arrow going to each vertex. Therefore our $\mathcal{R}$ is defined everywhere, but is not onto since 2 is not in the image (no arrow goes to 2).

REMARK 2.39. A play on words shows $\mathcal{R}$ is defined everywhere (resp. surjective) iff $\mathcal{R}^{-1}$ is surjective (resp. defined everywhere). The same duality holds between $\mathcal{R}$ being single-valued and injective.

2.3 Functions

While sets are the backbone of mathematics, functions are the tools which allow us to compare sets.

A function is a special kind of relation.

Definition 2.40. A relation $\mathcal{R}\colon A \rightsquigarrow B$ is a *function* (or mapping), denoted instead by $\mathcal{R}\colon A \to B$, if for every $a \in A$ there exists a unique $b \in B$ with $a \mathcal{R} b$. In this case we say that b is the image of a under $\mathcal{R}$.

We use the usual notation for a function.

Notation 2.41. For relations that happen to be functions, we use the more familiar notations $\mathcal{R}(a) = b$ or $\mathcal{R}\colon a \mapsto b$.

Functions will typically be denoted by the letters f, g, F, or G when they are between (plain) sets. When we start talking about groups, we will typically denote group homomorphisms (special kinds of functions between groups) by φ or ψ.

Another common notation for a function which records everything at once is used when the function is given by a concise rule. For instance,

$$\begin{aligned} F &: \mathbb{R} \to \mathbb{R} \\ &: x \mapsto x^3 + 1 \end{aligned}$$

is the function that takes a real number r and sends it to $r^3 + 1$. Of course in a basic math course this would be denoted as $F(x) = x^3 + 1$. Another example of this notation would be

$$\begin{aligned} F &: \mathbb{R}^2 \to \mathbb{R}^3 \\ &: (x, y) \mapsto \left(x^2 + x + y, x^4 + y^2, x^2 + y^3\right). \end{aligned}$$

For instance, with this function $F(1,1) = (3,2,2)$ and $F(1,2) = (4,5,9)$. Since functions are relations we have

Convention 2.42. Since a function is a special kind of relation, we can carry over all definitions from the theory of relations. For example:

(1) The image of a function $F\colon X \to Y$ is

$$\operatorname{Im}(F) \overset{\text{def}}{=} \{y \in Y : \exists x \in X \text{ s.t. } F(x) = y\}.$$

(2) A function $F\colon X \to Y$ is surjective (or onto) if $\operatorname{Im}(F) = Y$. In more detail

$$\forall y \in Y \, \exists x \in X \text{ s.t } F(x) = y.$$

(3) A function $F\colon X \to Y$ is injective (or one-to-one) if

$$\forall x, \dot{x} \in X, F(x) = F(\dot{x}) \Rightarrow x = \dot{x}.$$

In terms of functions, you should think of surjectivity as being able to solve any equation. For instance, asking whether the number 7 is in the image of the function

$$\begin{aligned} f &: \mathbb{R} \to \mathbb{R} \\ &: x \mapsto x^3 + 1 \end{aligned}$$

is the same as asking whether or not we can solve the equation

$$x^3 + 1 = 9.$$

Since $x = 2$ does indeed solve this equation, we have that $9 \in \mathrm{Im}(f)$. And asking whether or not this function is surjective is the same thing as asking if for all real number $r \in \mathbb{R}$ I can solve the equation

$$x^3 + 1 = r.$$

Again, we certainly can

$$x = \sqrt[3]{r - 1}$$

is a solution.

Recall that a relation $X \rightsquigarrow Y$ is defined everywhere if every $x \in X$ is related to at least one element of Y, and is single valued if every $x \in X$ is related to at most one element of Y. Thus a relation is defined everywhere and single valued iff every element is related to a unique element of Y, i.e. iff it is a function. Therefore we have the following characterization of functions:

A relation F, must satisfy the following two properties to be a function:

- F is defined everywhere.
- F is single-valued.

The first hypothesis that F be defined everywhere since in practice if we restrict any relation to its domain it becomes defined everywhere. Therefore we have the following

THEOREM 2.43. *Suppose that $F : X \rightsquigarrow Y$ is a single-valued relation. Then the restricted relation*

$$F|_D \stackrel{def}{=} F \circ \iota_D^X : D \rightsquigarrow Y$$

is a function iff $D \subseteq \mathrm{Dom}(F)$. ■

DEFINITION 2.44. We clearly have that ι_A^X is a function for any $A \subseteq X$ and the *restriction of $F : X \to Y$ to A*, denoted $F|_A$ is defined as $F \circ \iota_A^X$. It is a function from $A \to Y$. At the same time, if $F(X) \subseteq B \subseteq Y$, then the same relation defines a function $F : X \to B$. This is called the *corestriction of $f : X \to Y$ to B*. [I]

In less fancy language $F|_A$ is the function with domain A which corresponds to F. In calculus classes we restricted and corestricted functions without giving it much thought! For example, in such classes

$$\cos : \mathbb{R} \to \mathbb{R}, \quad \cos : \mathbb{R} \to [0, 1], \quad \text{and} \quad \cos : [0, 2\pi] \to [0, 1]$$

[I] We will not use separate notation for the corestriction of a function, but if we did $|_B F$ would be as good as any.

are all considered to be the same function, denoted simply as $\cos x$. This led to the very annoying definition of $\cos^{-1} x$ as the inverse function of $\cos x$, but only being on a small interval:

$$\cos^{-1} : [0,1] \to \left[0, \frac{\pi}{2}\right].$$

Of course we could have just as easily have taken the second set of the arccos to be $[-\pi/2, 0]$, or, if we're really feeling frisky, to be $[4\pi, 9\pi/2]$. In a calculus class domains where always taken to be maximal domains. Technically, a question which asks to find the domain of a function–which is a very common type of question in both calculus 1 and calculus 3– makes no sense since it is a part of the definition of a function. Really these types of questions are asking you to find the largest *possible* domain for a function.

For instance, in our definition

$$f(x) = \frac{1}{x-1}$$ [II]

is not a function $\mathbb{R} \to \mathbb{R}$ but it is a function $\mathbb{R} \setminus \{1\} \to \mathbb{R}$.

Remark 2.45. In this last example, the fact that 0 is not in the image is irrelevant. Notice that the relation f would also define a function $\mathbb{R} \setminus \{1\} \to \mathbb{R} \setminus \{0\}$.

In a way, the problem with the definition is that to us, the notion of a function makes sense without specifying it's first and second set. For instance, if I write the sentence "consider the function $f(x) = x^3 + 1$", this makes pretty good sense to us. However, to the math god, it makes zero sense because we have not specified first and second sets for it. For instance, $f(x) = x^3 + 1$ could have domain any set where addition and multiplication make sense. For instance, the domain could be $\mathbb{Z}_3$, in which case the function would be

$$f([0]_3) = [1]_3 \quad f([1]_3) = [2]_3 \quad f([2]_3) = [0]_3$$

which is a bijection. If the domain is $\mathbb{R}$, then basic calculus says that it's still a bijection. If we take the domain to be $\mathbb{Z}_7$, then the function [III] is

$$f([0]_3) = [1]_3 \quad f([1]_3) = [2]_3 \quad f([2]_3) = [0]_3$$

In general, in the context of a theory, two things in math are considered the same if there is no way of telling the different between them in the context of a theory.[IV] The natural definition of what it means for two functions, F and G, to be the same is that they are equal relations (i.e. they have the same graphs), but there is a slight problem with this. Namely that this ignores the second set which determines which functions can be composed with the original. For instance, the two functions

$$\begin{aligned} &\mathbb{R} \to [-1,1] && \mathbb{R} \to \mathbb{R} && (3)\\ &x \mapsto \cos x && x \mapsto \cos x && (4) \end{aligned}$$

have exactly the same graph: they are both

$$\{(x, \cos x) : x \in \mathbb{R}\}.$$

However, we can tell the difference between them in a practical way since the first function can be post-composed with the function

$$\frac{1}{2-x}$$

[II] In relation notation this says that

$$\mathcal{R} = \left\{(x,y) \in \mathbb{R} \times \mathbb{R} : y = \frac{1}{x-1}\right\}.$$

[III] Again, to the math god this makes no sense! We are not allowed to pick the function and then pick the domain!

[IV] Again, an example of this are congruent triangles in high school geometry.

whose first set does not contains 2, while the second one (whose second set is $\mathbb{R}$) cannot be. This annoyingly technical distinction is more a reflection of a flaw in set-theory than it is something we have to worry about for practical purposes. In fact we can replace the second set of both functions in Equation (3), by their images, and it completely disappears! However, it will turn out to be for the best to put it into our formal definition and then basically ignore it the second we leave this section! Therefore we give the following definition of what it means for two functions to be equal:

Definition 2.46. Two relations, F, G are said to be equal iff

- The first sets of both F and G are the same;
- If X denotes the common first set, and Y denotes this common second set, we require that the subset of $X \times Y$ they determine are also the same.

The following theorem translates this definition over to function notation. Its proof is just a simple exercise in definitions and is included in most (read all I've ever seen) foundations textbooks, c.f. [SSA01].

Theorem 2.47. *Let $F, G : X \to Y$ be two functions. Then $F = G$ iff*

$$\forall x \in X,\ F(x) = G(x).$$

■

In words, you can say that F and G are the same if they do the same thing to every point in their domain. When F and G are equal, we denote this as $F = G$. Note that with this definition, the two functions $F(x) = x^3 + x + 1$ and $G(t) = t^3 + t + 1$ are equal, but the functions $F(x) = 1$ and $G(x) = x/x$, having different domains, are not.

Theorem 2.48.

(1) *For any set X, $\mathbf{1}_X$ is a function.*

(2) *When defined, the composition of two function is a function.*

(3) *When defined, function composition is associative.*

(4) *For a function $f : X \to Y$,*

$$f \circ \mathbf{1}_X = f = \mathbf{1}_Y \circ f.$$

Proof. We have shown all but the second claim to be true for all relations. The second claim is true due to the first two parts of Exercise 2.26 on page 26. ■

Since if $X = Y$ every possible composition of functions is defined, thus, in this case, we can delete the words "when defined" in all instances of the above. In function notation, of course the identity function satisfies:

$$\mathbf{1}_X(x) = x$$

for all $x \in X$ and the composition of two functions f and g satisfy (and can be defined as)

$$(f \circ g)(x) = f(g(x)).$$

Noting this, proving that the composition of functions is associative is even easier than it was for relations, simply note that for all $x \in \mathrm{Dom}(h)$ that both $f \circ (g \circ h)$ and $(f \circ g) \circ h$ take x to $f\big(g(h(x))\big)$. and apply Theorem 2.47.

Warning! 2.49. What is the identity function in calculus? It certainly isn't the function $f(x) = 1$. (Don't confuse multiplication with composition!) This function is far from being injective, in fact it sends every single last real number to the number one, so it's infinite to 1. It's as far away from being injective as you can get. Of course, in reality, the identity function is the function $f(x) = x$ since for any number c, $f(c) = c$.

Remark 2.50. By "when defined" all I mean is that the first and second sets of the various functions line up. For instance, if I were to write out **(3)** out in detail, I would say suppose that $f: X \to Y$, $g: Y \to Z$ and $h: Z \to W$ are functions. Then

$$f \circ (g \circ h) \qquad (f \circ g) \circ h \qquad \text{are equal functions } X \to W.$$

Definition 2.51. A relation $F: X \rightsquigarrow Y$ is

(1) a *bijection* if it is a function and injective and surjective,

(2) is *invertible* if there exists a relation $G: Y \rightsquigarrow X$ which satisfies

$$F \circ G = \mathbf{1}_Y \quad \text{and} \quad G \circ F = \mathbf{1}_X. \tag{5}$$

Let $\mathcal{R}$ be a relation defined on a finite set X. A *path* in the digraph of $\mathcal{R}$ is exactly what you would it expect it to be, that it is it is a tuple

$$[a_1, \ldots, a_n]$$

with

$$(a_1, a_2),\ (a_2, a_3), \ldots,\ (a_{n-1}, a_n) \in \mathcal{R}.$$

We consider a path as an actual subset of the digraph which contains all of the relevant vertices and edges. We say that each a_i *lives* in $[a_1, \ldots, a_n]$ and $[a_1, \ldots, a_n]$ *contains* a_i. We call a path

$$[a_1, \ldots, a_n]$$

a *pre-circle* if exactly n points live in it (so for $1 \le i \ne j \le n, a_i \ne a_j$) and $(a_n, a_1) \in \mathcal{R}$. Notice that if $[a_1, \ldots, a_n]$ is a pre-circle circle, then so are any of its n cyclic permutations

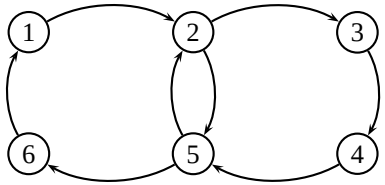
$$[a_1, \ldots, a_n],\ [a_n, a_1, a_2, \ldots, a_{n-1}],\ [a_{n-1}, a_n, a_1, \ldots, a_{n-2}], \ldots, [a_2, a_3, \ldots a_n, a_1].$$

Graphically, this corresponds to taking the same loop in the digraph of a relation but simply starting it from a different spot. We do not consider these pre-circles to be different. By the *circle*

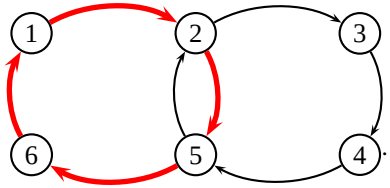
$$\langle\, a_1, \ldots, a_n \,\rangle$$

we mean any of the pre–circles above.

For instance,



contains the circles $\langle 1,2,5,6\rangle$, $G1,2,3,4,5,6$, $\langle 2,5\rangle$, and $\langle 2,3,4,5\rangle$. For instance, the circle $\langle 1,2,5,6\rangle$ is highlighted



Exercise 2.28.

(1) Define a relation on the set $\{1,2,3,4,5\}$ which has circles $\langle 1,2,4\rangle$, $\langle 2\rangle$, and $\langle 2,4,3,5\rangle$ and no others.

(2) Prove that $[a_1,\ldots,a_n]$ is a path in (the digraph of) $\mathcal{R}$ iff $[a_n,\ldots,a_1]$ is a path in (the digraph of) $\mathcal{R}^{-1}$.

(3) Prove that $\langle a_1,\ldots,a_n\rangle$ is a circle in (the digraph of) $\mathcal{R}$ iff $\langle a_n,\ldots,a_1\rangle$ is a circle in (the digraph of) $\mathcal{R}^{-1}$.

(4) Prove that $\mathcal{R}$ is a bijection iff the digraph of $\mathcal{R}$ can be partitioned into circle. (This means that there are circles $\mathbf{K}_1,\ldots,\mathbf{K}_m$ whose union is the entire digraph of $\mathcal{R}$ and $\mathbf{K}_i\cap\mathbf{K}_j=\varnothing$ whenever $1\le i\ne j\le m$.)

(5) Now we suppose further that $\mathcal{R}\colon X\rightsquigarrow X$ is single valued. Prove that $\mathcal{R}$ is a bijection iff every vertex of the digraph of $\mathcal{R}$ is contained in a unique circle.

Remark 2.52. To have everything in one place, let's record here what we need for a relation F to be a bijection.

- F is defined everywhere. That is
$$\forall x\in X,\exists y\in Y \text{ s.t. } (x,y)\in F.$$
- F is single valued. That is
$$\forall x,x'\in X,\forall y\in Y\left[(x,y)\in F \text{ and } (x',y)\in F\Rightarrow x=x'\right].$$
- F is onto (i.e. surjective). That is
$$\forall y\in Y,\exists x\in X \text{ s.t. } (x,y)\in F.$$
- F is one-to-one (i.e. injective). That is
$$\forall x\in X,\forall y,y'\in Y\left[(x,y)\in F \text{ and } (x,y')\in F\Rightarrow y=y'\right].$$

Note that the first two conditions, by definition, show us that F is a function and the second two conditions guarantee us that F^{-1} is a function (since the requirement that F^{-1} be single valued is exactly the same requirement that F be one-to-one, and the requirement that F^{-1} be defined everywhere is exactly the same requirement that F be onto). Thus, nearly by definition, we have that

$$F \text{ is a bijection } \Leftrightarrow F \text{ and } F^{-1} \text{ are functions.} \tag{6}$$

The following exercise was first mentioned in Rant 2.15. I hope it convinces you that the concept of a number is more basic than the concept of counting.

EXERCISE 2.29. All of my son's, who is a baby, outfits are snap ups. Explain in a mathematical way how I know that all of my son's outfits have as many "out" snaps as "in" buttons even though I have never counted them!

SOLUTION

The following theorem can be improved a bit by generalizing it to functions, but since we won't really need this we simply state and prove it for functions.

THEOREM 2.53. *For a function $F : X \to Y$, the following are equivalent*

(1) *F is a bijection.*

(2) *F is invertible.*

(3) *The relation F^{-1} is also a function.*

In which case we'll see in the proof $G = F^{-1}$ is the function which satisfies (5).

PROOF.

1 ⇔ 3 The conditions that F^{-1} is single-valued is the exactly the same condition that F be injective. The condition that F^{-1} is defined everywhere is exactly the same condition that F is surjective. Therefore the first and third statements are equivalent.

2 ⇒ 1 Now we show that if F is invertible then F is a bijection. Let G be a function which satisfies Equation (5). That is we are assuming that G satisfies

$$F \circ G = \mathbf{1}_Y \quad \text{and} \quad G \circ F = \mathbf{1}_X.$$

We need to show that F is a bijection, which by definition means that we need to show that F is injective and surjective. We start by showing that F is injective. Suppose that $F(x) = F(x')$ for elements $x, x' \in X$. We need to show that $x = x'$. But this follows from the calculation

$$x = \mathbf{1}_X(x) = G(F(x)) = G(F(x')) = \mathbf{1}_X(x') = x'.$$

Therefore F is injective.

We need to show that F is surjective. To that end let $y \in Y$ be arbitrary. We need to find (at least one) element of X which maps to y. I claim that $G(y)$ does just that. Indeed, if we denote $G(y)$ as x, then we have that

$$F(x) = F(G(y)) = \mathbf{1}_Y(y) = y.$$

Therefore F is surjective.

1 ⇒ 2 The assumption that F is bijective means that for every element of $y \in Y$ there exists a unique $x \in X$ so that $F(x) = y$. If we define $G(y)$ to be this element x, then G is a function. Moreover, by definition, we have that $F(G(y)) = y$, i.e., $FG = \mathbf{1}_Y$. The last thing we need to show is that $GF = \mathbf{1}_X$. But for any $x \in X$ we have that $G(F(x))$ is the unique element of X which F takes to $F(x)$. But x is such a point and since this point is unique we have that $G(F(x)) = x$ and therefore $GF = \mathbf{1}_X$ as desired. ■

Definition 2.54. For sets X and Y define the sets

(1) $\mathcal{F}\mathrm{un}(X,Y) \stackrel{\text{def}}{=} \{\varphi \mid \varphi\colon X \to Y\}$

(2) $\mathcal{F}_X \stackrel{\text{def}}{=} \{\varphi \mid \varphi\colon X \to X\} = \mathcal{F}\mathrm{un}(X,X)$

(3) $\mathcal{S}_X \stackrel{\text{def}}{=} \{\varphi \in \mathcal{F}_X \mid \varphi \text{ is a bijection }\} = \{\varphi \mid \varphi\colon X \to X, \varphi \text{ is a bijection }\}$.

A bijection whose first and second sets are the same is called a *permutation*. If these sets are both X (i.e. an element of $\mathcal{S}_X$), we say that the function is a *permutation* of X. Therefore, by definition

$$\boxed{\mathcal{S}_X = \text{Set of permutations of } X.}$$

For instance, the function $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 1$ defines a permutation of the set $\overline{3} = \{1,2,3\}$, and is hence an element of $\mathcal{S}_{\overline{3}}$, and $\tau(x) = y$, $\tau(y) = x$, $\tau(z) = z$ defines a permutation of the set $\{x,y,z\}$, and is hence an element of $\mathcal{S}_{\{1,2,3\}}$. In the next section we talk more about these, and we introduce some notation to describe them. We will spend a large part of the course studying these! The cool thing about these sets is that we don't have to worry about what we can and can't compose since for any two $f, g \in \mathcal{F}_X$, $f \circ g$ and $g \circ f$ make sense (the first set and the second set are the same!).

Theorem 2.55 (Pigeon Hole Principle). *Let X be a finite set and $\sigma\colon X \to X$ be a function. Then the following are equivalent:*

(1) *σ is a bijection,*

(2) *σ is an injection,*

(3) *σ is a surjection.*

Proof. {Supply or make an exercise} ∎

Exercise 2.30. Prove that

$$\#\left(\mathcal{F}\mathrm{un}(X,Y)\right) = \#Y^{\#X}.$$

Hint! Identify $\mathcal{F}\mathrm{un}(X,Y)$ with the cartesian product of $\#X$ copies of Y.

SOLUTION

Exercise 2.31. Let X be any set and let σ be any bijection from X to itself (i.e. a permutation of X).

(1) Suppose for some subset $A \subseteq X$ we have

$$\sigma(A) \subseteq A \quad \text{and} \quad \sigma^{-1}(A) \subseteq A. \tag{7}$$

Prove that $\sigma(A) = A$.

SOLUTION

(2) Prove that if A is finite, then we could replace (7) with just

$$\sigma(A) \subseteq A. \tag{8}$$

Hint! The restriction of an injection is an injection.

Of course, in general we really need both inclusions in Equation (7). For instance, if we consider the bijection (multiplication by 2)

$$\begin{aligned} m_2 &: \mathbb{Q} \to \mathbb{Q} \\ &: a \mapsto 2a \end{aligned}$$

then $m_2(\mathbb{Z}) \subseteq \mathbb{Z}$, but we do not have equality since $m_2(\mathbb{Z}) = 2\mathbb{Z}$, which does not include 1, but of course, $\mathbb{Z}$ does.

DEFINITION 2.56. Let $A \subseteq X$. A function $\mathrm{F}: X \to X$ is a *projection* onto A if $\mathrm{F} \circ \mathrm{F} = \mathrm{F}$. It is a *projection onto A* if the A is the image of F.

Of course every projection is a projection onto some set, namely its image.

EXERCISE 2.32. Show that a function $\mathrm{F}: X \to X$ is a projection onto A iff

SOLUTION

(1) For all $x \in X$, $\mathrm{F}(x) \in A$ ⟨Fancier: $\mathrm{F}(A) \subseteq A$. ⟩

(2) For all $a \in A$, $\mathrm{F}(a) = a$. ⟨Fancier: $\mathrm{F}|_A = \mathbf{1}_A$. ⟩

2.3.1 More on Permutations

Let's rerecord some of the facts about $\mathcal{S}_X$ for arbitrary sets X. Starting in Chapter **??**, these facts will be stated simply as "$\mathcal{S}_X$ is a group."

THEOREM 2.57.

Closure *If σ and τ are elements of $\mathcal{S}_X$, then so is $\sigma \circ \tau$.*[V]

Assoc. *For all σ, τ, and ρ in $\mathcal{S}_X$, we have that*

$$\sigma \circ (\tau \circ \rho) = (\sigma \circ \tau) \circ \rho.$$

Identity *The identity function, $\mathbf{1}_X$ is an element of $\mathcal{S}_X$. Moreover it satisfies*

$$\sigma \circ \mathbf{1}_X = \mathbf{1}_X \circ \sigma = \sigma$$

for all $\sigma \in \mathcal{S}_X$.

[V] In symbols, $\sigma, \tau \in \mathcal{S}_X \Rightarrow \sigma \circ \tau \in \mathcal{S}_X$

Inverses *For all $\sigma \in \mathcal{S}_X$ we have that σ^{-1} is also in $\mathcal{S}_X$ and*

$$\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \mathbf{1}_X.$$

Proof. We have already proven most of this already. For instance, associativity follows from the fact that the composition of relations is associative. Since permutations are special kinds of bijections, bijections are special kinds of functions, and functions are special kinds of relations this implies that the composition of permutations is associative. One way of showing that $\mathbf{1}_X$ is a bijection of X is noting that $\mathbf{1}_X = \mathbf{1}_X \circ \mathbf{1}_X$ implies that $\mathbf{1}_X$ is invertible with inverse equal to itself.[VI]

One sexy way of showing closure is as follows: Suppose that σ and τ are permutations of X. This means that σ^{-1} and τ^{-1} are functions. Moreover

$$\begin{aligned}
(\sigma \circ \tau) \circ (\tau^{-1} \circ \sigma^{-1}) &= \sigma \circ \left(\tau \circ (\tau^{-1} \circ \sigma^{-1})\right) && \text{(Associativity)} \\
&= \sigma \circ \left((\tau \circ \tau^{-1}) \circ \sigma^{-1}\right) && \text{(Associativity again)} \\
&= \sigma \circ (\mathbf{1}_X \circ \sigma^{-1}) && \text{(Since } \tau \circ \tau^{-1} = \mathbf{1}_X\text{)} \\
&= \sigma \circ \sigma^{-1} && \text{(Property of } \mathbf{1}_X\text{)} \\
&= \mathbf{1}_X. && \text{(Since } \sigma \circ \sigma^{-1} = \mathbf{1}_X\text{)}
\end{aligned}$$

[VII] Essentially the same argument proves that

$$(\tau^{-1} \circ \sigma^{-1}) \circ (\sigma \circ \tau) = \mathbf{1}_X$$

as well. This shows that $\sigma \circ \tau$ is invertible (with inverse $\tau^{-1} \circ \sigma^{-1}$). Therefore, Theorem 2.53 applies to show that $\sigma \circ \tau$ is also a bijection. ■

Really, the only place that we used the fact that the domain and codomain of a bijection were the same was in the fact that we composed the bijections without giving anything a second thought. Therefore the same proof will yield:

Theorem 2.58. *The composition (when defined) of any two bijections is a bijection. The inverse of any bijection is a bijection.* ■

In more detail, let X, Y, and Z be sets. Suppose that

$$\sigma \colon X \to Y \quad \text{and} \quad \tau \colon Y \to Z$$

are bijections. Then

$$\sigma^{-1} \colon Y \to X \quad \text{and} \quad \tau \circ \sigma \colon X \to Z$$

are bijections.

When considering the bijections of a set we typically study only the set

$$\overline{n} \stackrel{\text{def}}{=} \{1, \ldots, n\}$$

of positive integers between 1 and some positive integer n (this notation was first introduced in the introduction). For example,

$$\overline{4} = \{1, 2, 3, 4\}.$$

In this case we write $\mathcal{S}_{\overline{n}}$ simply as $\mathcal{S}_n$. For some strange reason, in this context we call the set $\overline{n}$ letters (instead of numbers). Thus, when we study say $\mathcal{S}_4$ we are studying the set of permutations on the 4 letters, $\{1, 2, 3, 4\}$. The advantage of studying $\mathcal{S}_n$ as opposed to $\mathcal{S}_X$ for some other set X with n things is that the set $\overline{n}$ comes to us already ordered with a nice, neat, pretty, red, bow. Therefore

> an element $\sigma \in \mathcal{S}_n$ can be thought of as a list of n things where the i^{th} entry expresses $\sigma(i)$. Moreover, in such a list every element of $\overline{n}$ appears once (rephrasing the fact that a bijection is onto) and only once (rephrasing the fact that a bijection is one-to-one).

[VI] Obviously the identity is injective and surjective, but the proof given makes no reference to elements of X and–as I hope I will convince you at some point–as such is superior.

[VII] These are functions, so I could have left the parenthesis out from the beginning, but it's good practice to see how this is implied by a simple associative property.

As such, we have

$$\mathcal{S}_n \overset{1\text{-}1}{\leftrightsquigarrow} \text{Lists of integers } 1 \leq i \leq n \text{ containing no repeats.}$$

Less importantly, but in a similar vein, we can think of $\mathcal{F}_n \overset{\text{def}}{=} \mathcal{F}_{\overline{n}}$ as

$$\mathcal{F}_n \overset{1\text{-}1}{\leftrightsquigarrow} \text{Lists of integers } 1 \leq i \leq n.$$

(Note that the set of lists of integers $1 \leq i \leq n$ are simply elements of the set $\overline{n}^n$ and lists containing no repeats are the elements of the set $\overline{n}^{*n}$ in our notation from earlier.) So the only difference is that the latter is allowed to repeat while the former is not.

To avoid confusion with future notation, and to make it easier on the eyes, we write these lists by writing the preimage above the image (we could, of course, just leave the top row off, but this would cause great confusion because we plan on using this notation later for something completely different when talking about a permutation) in the following way (for say $\sigma \in \mathcal{S}_6$) [VIII]

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \sigma(1) & \sigma(2) & \sigma(3) & \sigma(4) & \sigma(5) & \sigma(6) \end{pmatrix}.$$

For example, if we have that σ is the permutation

$$\sigma(1) = 2 \qquad \sigma(2) = 4 \qquad \sigma(3) = 6$$
$$\sigma(4) = 1 \qquad \sigma(5) = 3 \qquad \sigma(6) = 5$$

then in the new notation, we would write σ as

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 1 & 3 & 5 \end{pmatrix}.$$

Since the identity is the function $\mathbf{1}_{\overline{n}}$ with $\mathbf{1}_{\overline{n}}(i) = i$ for all $i \in \overline{n}$ in this notation, the identity permutation, $\mathbf{1}_{\overline{n}} \in \mathcal{S}_n$ is

$$\begin{pmatrix} 1 & 2 & 3 \ldots & n \\ 1 & 2 & 3 \ldots & n \end{pmatrix}$$

For instance, for $n = 3$, the identity is

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$$

and for $n = 6$ it's

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}.$$

Note also, in this notation, if we are given a permutation σ, we can find the inverse of σ by reading upwards instead of downwards (and then rearranging the top row to go in order). For example, if we are given that σ

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}$$

in $\mathcal{S}_4$, then σ^{-1} is the permutation

$$\begin{pmatrix} 3 & 2 & 4 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}.$$

[VIII] While these are written with parenthesis, these should in no way be thought of as matrices!

DEFINITION 2.59. A bijection from the set $\overline{n}$ to a set X with n elements is a *dictionary* on X or an *indexing* of X.

Notice that there are $n!$ dictionaries on a set with n elements. Notice that a dictionary is simply a way of counting a set. For instance, if I asked you how many elements are in the set $\{a, \spadesuit, \clubsuit, \blacklozenge, \blacktriangleright, 18, \pi, \diamondsuit\}$ you would stare at the set for a second and respond 8. How did you get that? Well, you probably looked at a and said 1, then $\spadesuit$ and said 2, etc. What you were really doing was looking at the bijection

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ a & \spadesuit & \clubsuit & \blacklozenge & \blacktriangleright & 18 & \pi & \diamondsuit \end{pmatrix}$$

which goes from $\overline{7}$ to the listed set.[IX] We will discuss this further in the optional section, Section 2.5 coming up on page 69 at the end of this chapter.

How many elements of $\mathcal{S}_n$ are there? Well, we can first specify where 1 goes. Then 2 can go to anything in $\overline{n}$ except for where 1 went. This leaves $n-1$ choices. Similarly, 3 can go to any of the $n-2$ places that 1 or 2 didn't go. In this way we see that $\mathcal{S}_n$ contains $n!$ (n factorial) elements, in total. We record this as a theorem.

THEOREM 2.60. $\#\mathcal{S}_n = n!$ ■

Interpreting $\mathcal{S}_n$ as elements of $\overline{n}^{*n}$, this is really a special case of Exercise 2.33. For instance, the proof in the special case that $n = 4$ is illustrated in the following diagram.

[IX] This way of explaining this is lifted nearly verbatim from [SSA01].

$1 \mapsto 1$	$2 \mapsto 2$	$3 \mapsto 3$	$4 \mapsto 4$
		$3 \mapsto 4$	$4 \mapsto 3$
	$2 \mapsto 3$	$3 \mapsto 2$	$4 \mapsto 4$
		$3 \mapsto 4$	$4 \mapsto 2$
	$2 \mapsto 4$	$3 \mapsto 2$	$4 \mapsto 3$
		$3 \mapsto 3$	$4 \mapsto 2$
$1 \mapsto 2$	$2 \mapsto 1$	$3 \mapsto 3$	$4 \mapsto 4$
		$3 \mapsto 4$	$4 \mapsto 3$
	$2 \mapsto 3$	$3 \mapsto 1$	$4 \mapsto 4$
		$3 \mapsto 4$	$4 \mapsto 1$
	$2 \mapsto 4$	$3 \mapsto 1$	$4 \mapsto 3$
		$3 \mapsto 3$	$4 \mapsto 1$
$1 \mapsto 3$	$2 \mapsto 1$	$3 \mapsto 2$	$4 \mapsto 4$
		$3 \mapsto 4$	$4 \mapsto 2$
	$2 \mapsto 2$	$3 \mapsto 1$	$4 \mapsto 4$
		$3 \mapsto 4$	$4 \mapsto 1$
	$2 \mapsto 4$	$3 \mapsto 2$	$4 \mapsto 1$
		$3 \mapsto 1$	$4 \mapsto 2$
$1 \mapsto 4$	$2 \mapsto 1$	$3 \mapsto 3$	$4 \mapsto 2$
		$3 \mapsto 2$	$4 \mapsto 3$
	$2 \mapsto 2$	$3 \mapsto 1$	$4 \mapsto 3$
		$3 \mapsto 3$	$4 \mapsto 1$
	$2 \mapsto 3$	$3 \mapsto 1$	$4 \mapsto 2$
		$3 \mapsto 2$	$4 \mapsto 1$

EXERCISE 2.33. Suppose that X is a set with n elements in it.

SOLUTION

(1) Prove for $m \leq n$ we have that

$$\#X^{*m} = \frac{n!}{(n-m)!}.$$

(2) Explain what this has to do with the fact that $\mathcal{S}_n$ has $n!$ elements that we just explained (the proof should almost "feel" the same).

We can use the same notations to denote a permutation of any set, however if the set is not a set of integers there is not natural choice of how to order the top row. For instance,

$$\begin{pmatrix} c & a & r & e & t \\ t & r & a & c & e \end{pmatrix}$$

denotes a permutation of the set of letters $\{a, c, e, r, t\}$ [x] which, for instance, takes c to t. Showing off (ok, I stole this!) we also have the bijection

$$\begin{pmatrix} a & h & i & p & g & u & y & s & t & o & r & e \\ h & e & y & i & t & s & a & g & r & o & u & p \end{pmatrix}.$$

We could even use the same notation for a bijection between two different sets or even any function between any two sets. (Of course the sets better be finite!) For instance,

$$\begin{pmatrix} \heartsuit & \clubsuit & \diamond & \spadesuit \\ \blacklozenge & \blacktriangleright & \bigstar & \blacksquare \end{pmatrix}$$

denotes the bijection from $\{\heartsuit, \clubsuit, \diamond, \spadesuit\}$ to $\{\blacklozenge, \blacktriangleright, \bigstar, \blacksquare\}$ which takes $\heartsuit$ to $\blacklozenge$, $\clubsuit$ to $\blacktriangleright$, $\diamond$ to $\bigstar$, and $\spadesuit$ to $\blacksquare$.

$$\begin{pmatrix} \heartsuit & \clubsuit & \diamond & \spadesuit \\ \blacklozenge & \blacktriangleright & \bigstar & \blacksquare \end{pmatrix}$$

Yet another way to represent a permutation is with a *braid diagram*. For a given permutation $\sigma \in \mathcal{S}_n$ this is simply a diagram with the numbers 1 through n on the top and bottoms and lines connecting the top number with the bottom number if σ of the top number is equal to the bottom number. Therefore it is nothing more than the digraph of the function from $\overline{n} \to \overline{n}$ (not making the usual convention that we have been making when the first set and second sets of a function are the same) drawn in a particular way.

DEFINITION 2.61. For a permutation σ we define

(1) The *crossing number* to be the number of crossings in the braid diagram of σ. In symbols it is

$$\text{Cross}(\sigma) \stackrel{\text{def}}{=} \#\{(i, j) : i < j, \sigma(i) > \sigma(j)\}.$$

(2) The *sign* of a permutation, denoted $\varepsilon(\sigma)$, is defined by the formula $(-1)^{Cross(\sigma)}$.

(3) A permutation is said to be *even* (and have even parity) if its crossing number is even (i.e. if its sign is $+1$) and *odd* (and have odd parity) if its crossing number is odd (i.e. if it's sign is -1).

(4) The set of all even permutations in $\mathcal{S}_n$ is denoted as $\mathcal{A}_n$. It is known as the *alternating group* on n-letters.

REMARK 2.62. If $\sigma \in \mathcal{F}_n$ (i.e. a function $\sigma : \overline{n} \to \overline{n}$) and not a bijection, we do not define its crossing number, but we do define the ε of such a function to always be 0, i.e. $\varepsilon(\sigma) = 0$ for all functions which are not bijections (permutations).

That is the crossing number of σ is the number of pairs of letters whose order gets flipped around after applying σ. We will later give a different definition of the sign of a permutation, and then show the two definitions given are the same. There is no notation for the set of odd permutations. While the crossing number provides more detailed information, it does not have nearly as nice of properties (see Exercise 2.39 below).

[x] o.k there is a usual order (albeit totally arbitrary) on the set of letters

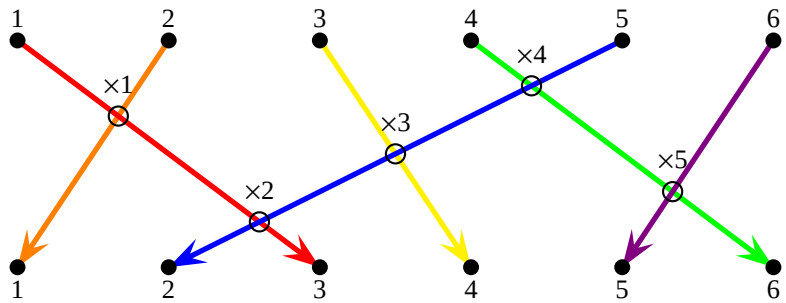
Remark 2.63. We give these definitions now because they are probably the simplest definition (i.e. it is the intuitive definition) of the *parity* of a permutation. However, later in Definition 8.31 we will redefine these in a way which will make it obvious that, for instance, an even permutation composed with an even permutation is even. It will also become clear with this new definition that 1/2 of all permutations are even (and hence half are also odd). For example, $\mathcal{A}_4$ has $24/2 = 12$ elements in it.

Example 2.64.

(1) The braid diagram for the permutation [XI]

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 4 & 6 & 2 & 5 \end{pmatrix}$$

is



We can see from the diagram that this permutation has crossing number 5. Indeed the 5 crossings are labeled. More explicitly the pairs which cross are

$$(1,2) \qquad (1,5) \qquad (3,5) \qquad (4,5) \qquad (4,6).$$

Indeed we have that

$$\begin{array}{ccc} 1 & < & 2 \\ \downarrow & & \downarrow \\ 3 & > & 1 \end{array} \qquad \begin{array}{ccc} 1 & < & 5 \\ \downarrow & & \downarrow \\ 3 & > & 2 \end{array} \qquad \begin{array}{ccc} 3 & < & 5 \\ \downarrow & & \downarrow \\ 4 & > & 2 \end{array} \qquad \begin{array}{ccc} 4 & < & 5 \\ \downarrow & & \downarrow \\ 6 & > & 2 \end{array} \qquad \begin{array}{ccc} 4 & < & 6 \\ \downarrow & & \downarrow \\ 6 & > & 5 \end{array}$$

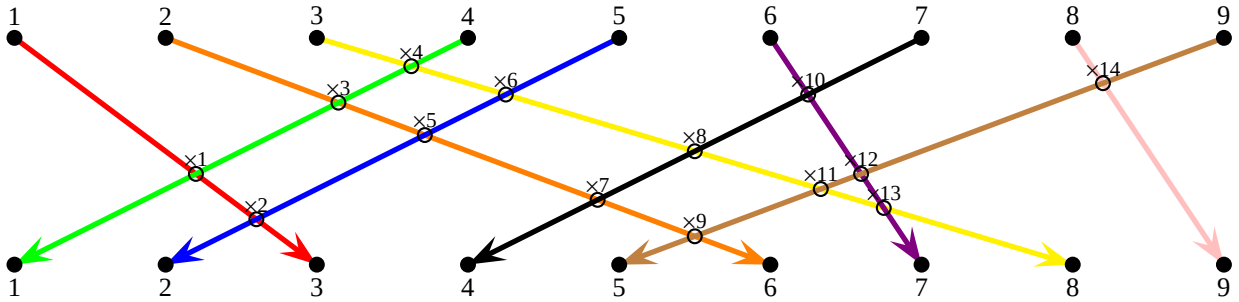
Since 5 is odd the permutation is also odd and it is **not** an element of $\mathcal{A}_6$.

(2) The braid diagram for the permutation

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

is

[XI] In these examples I'm using colors only because I feel as if though they make it easier to see where the crossings occur. They have no specific meaning.



Therefore the crossing number of this permutation is 14, the permutation is even and, by definition, is an element of $\mathcal{A}_9$. It should be noted that actually drawing this diagram with your free hand (instead of using the computer which draws everything perfectly) could be rather annoying! (Try it without looking here for help!)

EXERCISE 2.34. Calculate the crossing number and parity of the following permutations:[XII]

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 1 & 3 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 6 & 2 & 4 & 5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 4 & 8 & 1 & 2 & 7 & 5 & 9 & 6 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 1 \end{pmatrix}$$

SOLUTION

EXERCISE 2.35. Prove

$$\sigma = \mathbf{1} \Leftrightarrow \text{Cross}(\sigma) = 0.$$

Careful! Obviously you need to use the formula since you can't draw a general enough picture.

SOLUTION

EXERCISE 2.36. Show that a permutation and its inverse have the same crossing number. Conclude that they also have the same parity.
Careful! Obviously you need to use the formula since you can't draw a general enough picture to give a proof! (This shouldn't stop you from drawing your own pictures on scrap paper to build your intuition enough to be able to come up with an abstract proof!)

SOLUTION

EXERCISE 2.37. Find an element of $\mathcal{S}_4$ whose crossing number is

(1) 0 **(2)** 1 **(3)** 2 **(4)** 3 **(5)** 4 **(6)** 5 **(7)** 6

SOLUTION

Exercise 2.38. Prove that

$$\binom{n}{2} = \max_{\sigma \in \mathcal{S}_n} \text{Cross}(\sigma).$$[XIII]

Solution

Hint! Consider a permutation whose braid diagram is a 'half twist.'
Hint! Don't forget that

$$\binom{n}{2} = 1 + 2 + 3 \cdots (n-2) + (n-1).$$

Exercise 2.39. Find permutations σ and τ with

Solution

(1) $\text{Cross}(\sigma \circ \tau) \neq \text{Cross}(\sigma) + \text{Cross}(\tau)$,

(2) $\text{Cross}(\sigma \circ \tau) \neq \text{Cross}(\sigma) \cdot \text{Cross}(\tau)$.

The next exercise is a sort of triangle inequality for the crossing number.

Exercise 2.40. Prove for any $\sigma, \tau \in \mathcal{S}_n$ we have

Solution

$$\text{Cross}(\sigma \circ \tau) \leq \text{Cross}(\sigma) + \text{Cross}(\tau).$$

Definition 2.65. Suppose that $f : X \to Y$ is a function and $A \subseteq X$ and $B \subseteq Y$ define $f(A)$ and $f^{-1}(B)$ as they were defined for relations. More consciously:

(1) $f(A) \stackrel{\text{def}}{=} \{f(a) \in Y : a \in A\}$

(2) $f^{-1}(B) \stackrel{\text{def}}{=} \{x \in X : f(x) \in B\}$.

In this situation $f(A)$ is called *the image of A under f* and $f^{-1}(B)$ is called *the inverse image of B under f*. We simply call the image of X under f, denoted $\text{Im}(f)$, *the image of f*. Note that by definition of a function (since a function is defined everywhere), $f^{-1}(\text{Im}(f)) = X$. We should think of $f(A)$ as $a \in A \Rightarrow f(a) \in f(A)$ and $f^{-1}(B)$ as saying $x \in f^{-1}(B) \Leftrightarrow f(x) \in B$. For instance, if $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x, y) = x^2 + y^2$, then $f^{-1}([0, 1))$ is the (interior of) the unit disc

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$$

and the image of the function is all positive real numbers.

Theorem 2.66.

(1) *$f(f^{-1}(B)) \subseteq B$ and equality holds if f is surjective.*

(2) *$A \subseteq f^{-1}(f(A))$ and equality holds if f is injective.*

Proof.

(1) Let $y \in f(f^{-1}(B))$ be arbitrary. Then there is an $x \in f^{-1}(B)$ so that $f(x) = y$. But since $x \in f^{-1}(B)$ we must have that $f(x) \in B$. But $y = f(x)$ so $y \in B$. Since y was arbitrary, $f(f^{-1}(B)) \subseteq B$.

Now suppose that f is surjective and let $b \in B$ be arbitrary. Since f is surjective, there exists an $x \in X$ so that $f(x) = b$. Since $f(x) \in B$ we have that $x \in f^{-1}(B)$ and so $f(x) \in f(f^{-1}(B))$. But $f(x) = b$ so $b \in f(f^{-1}(B)$. Since b was arbitrary, we have that $B \subseteq f(f^{-1}(B))$

(2) Let $a \in A$ be arbitrary. To show that $a \in f^{-1}(f(A))$ we need to show that $f(a) \in f(A)$, but this is true by the definition of $f(A)$. Since a was arbitrary $A \subseteq f^{-1}(f(A))$.

Suppose now that f is one-to-one and let $x \in f^{-1}(f(A))$ be arbitrary.

$$\begin{aligned}
&\Rightarrow f(x) \in f(A) && \text{(defn of } f^{-1}\text{)}\\
&\Rightarrow \exists a \in A \text{ s.t. } f(a) = f(x) && \text{(defn of } f(A)\text{)}\\
&\Rightarrow a = x && (f \text{ is one to one})\\
&\Rightarrow x \in A && \text{(since } x = a \in A\text{).}
\end{aligned}$$

Since x was arbitrary, we have shown that $f^{-1}(f(A)) \subseteq A$. ■

Exercise 2.41. The following statement (and hence its proof) is false. Find the error in the proof.

False Theorem 2.67. *For any function f we have that $A = f^{-1}(f(A))$.*

"**Proof**". We have already seen that the inclusion, $A \subseteq f^{-1}(f(A))$, holds. Therefore, we only need to prove that the other inclusion is valid. To this end let x be in the set on the right hand side, $f^{-1}(f(A))$. Then $y = f(x) \in f(A)$. But this means that there exists an $x \in A$ so that $f(x) = y$. Therefore $x \in A$, and we have proven the reverse inclusion. ■

SOLUTION

Neither of these inequalities are equalities if the extra stated conditions does not hold. For instance, consider the constant function from $\mathbb{R}$ to itself defined by $f(x) = 5$. Then

$$f(\{0\}) = \{5\} \Rightarrow f^{-1}(f(\{0\})) = \mathbb{R}$$

and

$$f^{-1}([3,6]) = \mathbb{R} \Rightarrow f(f^{-1}([3,6])) = \{5\}.$$

Exercise 2.42. Suppose that $f : X \to Y$ is a function and $A_1, A_2 \subseteq X$ and $B_1, B_2 \subseteq Y$. Then prove

(1) $A_1 \subseteq A_2 \Rightarrow f(A_1) \subseteq f(A_2)$,

(2) $B_1 \subseteq B_2 \Rightarrow f^{-1}(B_1) \subseteq f^{-1}(B_2)$,

(3) $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$,

(4) $f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$,

SOLUTION

(5) $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$,

(6) $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$,

(7) $f^{-1}(B_1 \setminus B_2) = f^{-1}(B_1) \setminus f^{-1}(B_2)$,

(8) $f|_{A_1}$ is a function from $A_1 \to B_1 \Leftrightarrow f(A_1) \subseteq B_1$.

Exercise 2.43. (**The Vertical Line Test**) Suppose that $\mathcal{R}\colon X \rightsquigarrow Y$ is a relation.

Solution

(1) Prove that

$$\begin{aligned} \pi_{\mathcal{R}}\colon \mathcal{R} &\to X && (9)\\ \colon (x,y) &\mapsto x && (10)\end{aligned}$$

is a function.

(2) Suppose that the relation $\mathcal{R}$ is a function. Prove that $\pi_{\mathcal{R}}$ is a bijection.

(3) Conversely, prove that if $\pi_{\mathcal{R}}$ is a bijection, then $\mathcal{R}$ is a function.

Exercise 2.44. Let $A \subseteq X$. Define $\Delta(A) := \{(x,x) : x \in A\}$. Prove that

Solution

$$\begin{aligned} &\colon A \to \Delta(A)\\ &\colon a \mapsto (a,a)\end{aligned}$$

is a bijection.

Exercise 2.45. Suppose that

Solution

$$f\colon X \to Z \quad \text{and} \quad g\colon Y \to W$$

are functions.

(1) Prove that

$$\begin{aligned} (f,g)\colon X \times Y &\to Z \times W\\ \colon (x,y) &\mapsto (f(x), g(y))\end{aligned}$$

is a function.

(2) Prove that (f, g) is a bijection iff f and g are bijections.

(3) Suppose that $X = Y$, $Z = W$ and $f = g := F$. Prove that for any relation $\mathcal{R}$ defined on X that

$$(F, F)(\mathcal{R}) = F \circ \mathcal{R} \circ F^{-1}.$$

(i.e. both are the same subset of $Y \times Y$.)

2.3.2 What is Calculus Saying?*

In this section we take a brief look at some functions from $\mathbb{R} \to \mathbb{R}$ and describe how the concepts of these sections apply by using some Calculus I stuff.

Example 2.68. Let

$$f(x) = \frac{27}{32}\left(x^3 + x^2 - x + \frac{5}{27}\right).$$

I claim that $f(x)$ defines a surjective function from $[0, 1] \to [0, 1]$. That is we need to show that $f([0, 1]) \subseteq [0, 1]$. By the extreme value theorem, this function has a maximum and a minimum when restricted to the (closed and bounded) interval $[0, 1]$. If this max/min is on $(0, 1)$, then it must be a local max/min and the derivative must vanish there. Now

$$f'(x) = \frac{27}{32}\left(3x^2 + 2x - 1\right) = 0 \Leftrightarrow x \in \left\{\frac{1}{3}, -1\right\}.$$

Hence the maximum and the minimum of the function restricted to $[0, 1]$ must occur at $x = 1/3$, 0, or 1. Now,

$$f(0) = \frac{5}{27}, \qquad f\left(\frac{1}{3}\right) = 0, \qquad f(1) = 1.$$

Hence the max of the function on $[0, 1]$ is 1 and the min of the function is 0, and we get $f([0, 1]) \subseteq [0, 1]$ as desired. The intermediate value theorem then implies the function is surjective. However, the function is not one to one since the function must take all values in $\left[0, \frac{5}{27}\right]$ at least twice (once before $1/3$ and once after.)

2.4 Equivalence Relations

In everyday life the notion of equality is a vague concept and often depends on context. If a 10 year old boy with chalk and an 87 year old martian with red pen each write an "x" somewhere, they have both written the same letter. However, they didn't write the same "'x". Life is full of these little nuances. As a kid I was devastated to learn that protestors had burned *the* American flag. However, my parents patiently explained to me that there really is no such thing as *the* American flag. Thus at the same time everybody who owns that red white and blue stared and striped fabric owns the American flag and nobody at all owns it. This same things applies for in an uncountable number of situations. The English language is full ways one can confuse the articles "the" and "a". Math is not much different. For instance, is the number $2/4$ the same as the number $1/2$? If they were really identical, why do they have different numerators? Are the expressions

$$\frac{d}{dx}\sin(x) \quad \text{and} \quad \cos(x)$$

the same? If so, why would would a student get points taken off on a Calculus I exam for answering a derivative question with the first answer but not the second? Are the functions $\frac{x}{x}$ and 1 the same? Why can't you evaluate the first one at zero? How about $\frac{x^2+1}{x^2+1}$ and 1? The thing is that in mathematics **equality depends on context.** For example, we have already discussed how technically

$$(A \times B) \times C \neq A \times (B \times C) \neq A \times B \times C$$

unless one of the sets A, B or C are empty. (This was because the ordered pair $((a,b),c)$ is not the same as the ordered pair $(a,(b,c))$ since $(a,b) \neq a$ and (only need or) $c \neq (b,c)$.) However this distinction is again superficial in most contexts of mathematics (i.e. all that I have ever seen and ever hope to see!).

Another example is the distinction between sets which contains a single object, such as $\{a\}$, and the object a itself.

Definition 2.69. Let $\mathcal{R}$ be a relation on a set A. Then $\mathcal{R}$ is

- $\mathcal{R}$ is *reflexive* if $a \in A \Rightarrow a\,\mathcal{R}\,a$.
- $\mathcal{R}$ is *symmetric* if $a\,\mathcal{R}\,b \Rightarrow b\,\mathcal{R}\,a$.
- $\mathcal{R}$ is *anti–symmetric* if $a\,\mathcal{R}\,b$ and $b\,\mathcal{R}\,a$ implies that $a = b$.
- $\mathcal{R}$ is *transitive* if $a\,\mathcal{R}\,b$ and $b\,\mathcal{R}\,c$ implies that $a\,\mathcal{R}\,c$.

Definition 2.70. Let $\mathcal{R}$ be a relation on a set A. Then $\mathcal{R}$ is

(1) an *equivalence relation* if it is

- **(a)** reflexive,
- **(b)** symmetric, and
- **(c)** transitive.

(2) a *partial order* if it is

- **(a)** reflexive,
- **(b)** anti–symmetric, and
- **(c)** transitive.

Superficially, the only difference between an equivalence relation and a partial order is the insertion of the word "anti" in the second part of the definition.

NOTATION 2.71. For an equivalence relation, we call related elements a and b *equivalent* and if $a\mathcal{R}b$ for a partial order we say that a is *smaller than or equal to* b.

Typically, equivalence relations are denoted $\sim$ and partial orders are denoted $\preccurlyeq$ instead of the symbol $\mathcal{R}$ which is typically only used for generic relations. In our course, equivalence relations are far more important than posets. When creating the digraph of a posets one typically only puts arrows connecting immediate successors (i.e. c is an *immediate successor* of a if we have that $a \preccurlyeq b \preccurlyeq c$ implies that $a = b$ or $b = c$.)

EXAMPLE 2.72.

(1) For any set X, $\mathbf{1}_X$ is an equivalence relation on X. In words it is the relation that x and y are related iff they are equal. Clearly it is the smallest equivalence relation one can define on a set.

(2) For any set X, the relation $X \times X$ is an equivalence relation. In this equivalence relation any two elements of X are related. Clearly it is the largest equivalence relation one can define on a set. Note that since every relation defined on X is by definition a subset of $X \times X$ every relation is contained in some equivalence relation.

(3) The relation $a \leq b$ is a partial order on $\mathbb{R}$ (since $a \leq a$, $a \leq b$ and $b \leq a$ implies that $a = b$ and $a \leq b$ while $b \leq c$ implies that $a \leq c$.) Similarly, the relation $a \geq b$ is also a partial order.

(4) The relation on the collection of all sets given by $X \subseteq Y$ is a partial order.

(5) The relation on the collection of all sets given by $X = Y$ is an equivalence relation.

(6) As usual $\mathbb{Z}$ is the set of integers. Let a and $b \in \mathbb{Z}$. We say that a *divides* b or that b is a *multiple* of a if there exists an integer m so that $b = ma$. This is denoted as $a \mid b$ and is considered a relation on the integers. The relation is not quite antisymmetric on the set of *all* integers since for every integer a, $a \mid (-a)$ and $(-a) \mid a$, but $a \neq -a$ unless $a = 0$. However, this relation does in fact define a partial order if we restrict attention to only the set of positive integers, $\mathbb{Z}^+$ (or just the set of negative integers), since

Reflexive For any $a \in \mathbb{Z}^+$ we have that $a = 1a$ which implies that $a \mid a$ (since there exists an integer n so that $a = na$, namely $n = 1$).

Anti-Symmetric Suppose that $a \mid b$ and $b \mid a$. There there are integers m_1 and m_2 with $b = m_1 a$ and $a = m_2 b$. This implies that $a = m_1 m_2 a$ so that, after canceling a from both sides, $m_1 m_2 = 1$. This can only happen if m_1 and m_2 are both equal to 1 or to -1. But the latter case implies that $b = -a$ which is impossible since both a and b are assumed to be positive. Therefore $a = b$.

Transitive Suppose that $a \mid b$ and $b \mid c$. Then there are integers m and n so that $b = na$ and $c = mb$. Thus

$$c = mb = m(na) = (mn)a = Na$$

where N is an integer (which happens to be equal to $m \cdot n$). That is $a \mid c$.

(7) The relation $\mathfrak{p}$ defined on the set of all people, $\mathcal{P}$, given by $(a, b) \in \mathfrak{p}$ iff a and b are the same or siblings is an equivalence relation on $\mathcal{P}$ which is the same relation relating people who have the same parents.

(8) If f is any function between any two sets, then define the relation $\sim_f$ via

$$a \sim_f b \Leftrightarrow f(a) = f(b).$$

This relation is an equivalence relation (defined on the domain of the function f). Indeed

Reflexive For any point a, $f(a) = f(a)$, implying that $a \sim_f a$ (so $\sim_f$ is reflexive.)

Symmetric If $a \sim_f b$, then $f(a) = f(b)$. This means that $f(b) = f(a)$ which implies that $b \sim_f a$.

Transitive Suppose that $a \sim_f b$ and $b \sim_f c$. Then $f(a) = f(b)$ and $f(b) = f(c)$. This implies that $f(a) = f(c)$ and so $a \sim_f c$.

EXERCISE 2.46. Did we use the fact that f was a function (and not just a relation) anywhere in this proof? If so where?

SOLUTION

EXAMPLE 2.73.

(1) Define a relation $\asymp$ on the set of non-zero elements of the set of non-zero n-vectors, $\mathbb{R}^n \setminus \mathbf{0}$, via $\mathbf{v} \asymp \mathbf{w}$ whenever there is a non-zero $k \in \mathbb{R}$ so that $k \cdot \mathbf{v} = \mathbf{w}$. Two related vectors are said to be *parallel*.[1] Then $\asymp$ is an equivalence relation since:

Reflexive Let $\mathbf{v} \in \mathbb{R}^n$. Then $k = 1$ is a non-zero real number with $1 \cdot \mathbf{v} = \mathbf{v}$. Hence $\mathbf{v} \asymp \mathbf{v}$ (via 1) and $\asymp$ is reflexive.

Symmetric Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ with $\mathbf{v} \asymp \mathbf{w}$. Then there is some non-zero $k \in \mathbb{R}$ with $k \cdot \mathbf{v} = \mathbf{w}$. But then k^{-1} is also some non-zero real number and $k^{-1} \cdot \mathbf{w} = \mathbf{v}$. Therefore $\mathbf{w} \asymp \mathbf{v}$ (via k^{-1}).

Transitive Let $\mathbf{v}, \mathbf{w}, \mathbf{u} \in \mathbb{R}^n$ with $\mathbf{v} \asymp \mathbf{w}$ and $\mathbf{w} \asymp \mathbf{u}$. Then there are non-zero real numbers k_1 and k_2 with $k_1 \cdot \mathbf{v} = \mathbf{w}$ and $k_2 \cdot \mathbf{w} = \mathbf{u}$. Then $k := k_2 k_1$ is a non-zero real number, and we have

$$k \cdot \mathbf{v} = (k_2 k_1) \cdot \mathbf{v} = k_2 \cdot (k_1 \cdot \mathbf{v}) = k_2 \cdot \mathbf{w} = \mathbf{u}.$$

Hence $\mathbf{v} \asymp \mathbf{u}$ (via k).

Geometrically, two points non-zero points are related to each other under this relation iff they are contained in the same line through the origin.

(2) The following relation will be one of the most important equivalence relations studied in this class. Let n be a positive integer. On the set $\mathbb{Z}$ of integers, we define a relation $\equiv_n$ via

$$a \equiv_n b \Leftrightarrow n \mid (a - b).$$

This is an equivalence relation since

Reflexive $a \equiv_n a$ since $0 \cdot n = a - a$.

Symmetric If $a \equiv_n b$, then there exists an integer m so that $mn = a - b$. But then $-m$ is also an integer and $(-m)n = b - a$. That is $n \mid (b - a)$ and so $b \equiv_n a$.

Transitive If $a \equiv_n b$ and $b \equiv_n c$, then there are integers m_1 and m_2 so that $m_1 n = a - b$ and $m_2 n = b - c$. Now let $m = m_1 + m_2$. Since the sum of any two integers is an integer, we have that $m \in \mathbb{Z}$. Adding these equations together shows us that

$$mn = (m_1 + m_2)n = m_1 n + m_2 n = (a - b) + (b - c) = a - c.$$

Hence $n \mid (a - c)$ and $a \equiv_n c$.

Therefore, for any integer n, we have an equivalence relation, $\equiv_n$!

[1] Later, we will consider the zero vector, $\mathbf{0}$, to be parallel to all other vectors. However if this were our relation, it would no longer be an equivalence relation since transitivity would fail.

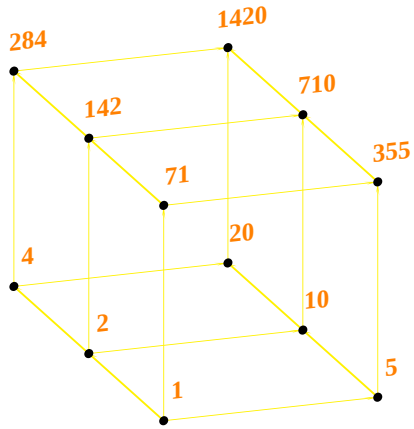


Figure 2.5: Lattice of Divisors of 1420.

EXERCISE 2.47. Show that $a \equiv_2 b$ iff a and b have the same parity (i.e. are both even or both odd).

SOLUTION

EXERCISE 2.48. Prove that the lattice of (positive) divisors of a positive integer looks the same whether viewed from top to bottom or bottom to top.
Hint! The hardest part of this problem is making more precise the notion of "looking the same." If you find yourself "talking" too much in your reasoning your are probably doing the problem wrong!

SOLUTION

EXERCISE 2.49. Let $X = \mathbb{Z} \times \mathbb{Z}^\times$. Define a relation $\sim$ on X by $(a, b) \sim (c, d)$ iff $ad - bc = 0$.

(1) Show that $\sim$ is an equivalence relation.
Bonus! Do this without dividing.

(2) Show that if we defined the same relation $\sim$ but instead we considered X as $\mathbb{Z} \times \mathbb{N}$, then we would no longer have an equivalence relation.

SOLUTION

EXERCISE 2.50. Suppose that $\sim$ is an equivalence relation defined on a set X and that σ is a permutation of X. Prove that $(\sigma, \sigma)(\sim)$ is also an equivalence relation on X.

SOLUTION

Hint! Unwinding the notation, if we denote the new relation

$$\sim' := (\sigma, \sigma)(\sim),$$

then we have that $a \sim b$ iff $\sigma(a) \sim' \sigma(b)$.

EXERCISE 2.51. Suppose that $\sim$ and $\sim'$ are equivalence relations on a set X. Prove that

SOLUTION

(1) $\sim \cap \sim'$ is an equivalence relation.

(2) $\sim \circ \sim' \cup \sim' \circ \sim$ is an equivalence relation.

Prove that the intersection of any collection of equivalence relations is an equivalence relation.

EXERCISE 2.52. Show for any relation $\mathcal{R}$ on A there exists a unique smallest equivalence relation $\widetilde{\mathcal{R}}$ which contains $\mathcal{R}$. (This means if $\sim$ is an equivalence relation containing $\mathcal{R}$, then $\widetilde{\mathcal{R}} \subseteq \sim$).

SOLUTION

Hint! Consider the intersection of all equivalence relations containing $\mathcal{R}$. (Why is this collection non–empty?)

The equivalence relation $\widetilde{\mathcal{R}}$ defined in the above exercise is called the equivalence relation *determined by* or *generated by* by $\mathcal{R}$. It has the following three properties (which we will see quite a few times!)

(1) $\mathcal{R} \subseteq \widetilde{\mathcal{R}}$,

(2) $\widetilde{\mathcal{R}}$ is an equivalence relation,

(3) If $\sim$ is an equivalence relation and $\mathcal{R} \subseteq \sim$, then $\widetilde{\mathcal{R}} \subseteq \sim$.

DEFINITION 2.74. We say that a permutation $\sigma\colon X \to X$ preserves a relation, $\mathcal{R}$, defined on X if $(\sigma, \sigma)(\mathcal{R}) = \mathcal{R}$.

EXERCISE 2.53. Suppose that σ is a permutation of X and that $\mathcal{R}$ is a relation defined on X. Prove that if σ preserves $\mathcal{R}$, then σ also preserves $\widetilde{\mathcal{R}}$.

SOLUTION

Exercise 2.54. Let

$$\mathcal{R}_1 := \{(a,b) \in \mathbb{R}\times\mathbb{R} : ab = a+1\} \qquad \mathcal{R}_2 := \{(a,b) \in \mathbb{R}\times\mathbb{R} : a^2 = b^2\}$$
$$\mathcal{R}_3 := \{(a,b) \in \mathbb{R}\times\mathbb{R} : a^3 = b^2\} \qquad \mathcal{R}_4 := \{(a,b) \in \mathbb{R}\times\mathbb{R} : ab = 1\}$$
$$\mathcal{R}_5 := \{(a,b) \in \mathbb{R}\times\mathbb{R} : a^2+b^2 = -1\} \qquad \mathcal{R}_6 := \{(a,b) \in \mathbb{C}\times\mathbb{R} : a^2+b^2 = -1\}$$
$$\mathcal{R}_7 := \{(a,b) \in \mathbb{C}\times\mathbb{C} : a^2+b^2 = -1\} \qquad \mathcal{R}_8 := \{(a,b) \in \mathbb{Q}\times\mathbb{Z} : ab \in \mathbb{Z}\}$$
$\mathcal{R}_9$:= The first relation from Exercise 2.14 $\qquad$ $\mathcal{R}_{10}$:= The second relation from Exercise 2.14

(1) Calculate the domain and image of each relation. Try to make your answer as clear and concise as possible: for instance

$$\mathrm{Dom}(\mathcal{R}_6) = \{ri \in \mathbb{C} : r \in \mathbb{R}, |r| \geq 1\}$$

is a pretty concise answer. Can you do even better?

(2) Which relations are single-valued? Which are one-to-one?

(3) (Only applies for the relations whose first set is the same as their second) Which are symmetric? Which are reflexive? Which are transitive? Which are equivalence relations?

SOLUTION

Exercise 2.55. Show that an equivalence relation is defined everywhere.

SOLUTION

Exercise 2.56. Show that a relation $\mathcal{R}$ defined on X is

(1) reflexive iff $\mathbf{1}_X \subseteq \mathcal{R}$,

(2) symmetric iff $\mathcal{R}^{-1} = \mathcal{R}$,

(3) transitive iff $\mathcal{R} \circ \mathcal{R} = \mathcal{R}$.

SOLUTION

Exercise 2.57. Let $\mathcal{R}$ and $\mathcal{S}$ be relations. Prove that

(1) If $\mathcal{R}$ and $\mathcal{S}$ is reflexive, then so is $\mathcal{R} \cap \mathcal{S}$.

(2) If $\mathcal{R}$ and $\mathcal{S}$ are symmetric, then so is $\mathcal{R} \cap \mathcal{S}$.

(3) If $\mathcal{R}$ and $\mathcal{S}$ are transitive, then so is $\mathcal{R} \cap \mathcal{S}$.

(4) If $\mathcal{R}$ and $\mathcal{S}$ are functions, then $\mathcal{R} \cap \mathcal{S}$ is a function iff $\mathcal{R} = \mathcal{S}$.

SOLUTION

(5) If $\mathcal{R}$ and $\mathcal{S}$ are functions, then $\mathcal{R} \cup \mathcal{S}$ is a function iff $\mathcal{R} = \mathcal{S}$.

Exercise 2.58. Suppose that $\mathcal{R}$ is a relation which contains a function F. Show that $\mathcal{R}$ is a function iff $\mathcal{R} = F$.

Solution

Exercise 2.59. Show that a relation which is both reflexive and a function is the identity.

Solution

Definition 2.75. Let $(\mathcal{P}, \preccurlyeq)$ be a poset. An element $p \in \mathcal{P}$ is called *maximal* if $p \preccurlyeq p'$ implies that $p = p'$. A *minimal* element is defined similarly.[II]

In terms of the poset diagram, these are the elements on the top of the diagram. For instance, if we look at the poset of divisors of the number 1420 but not equaling 1420 or 1 we have the set

$$\{2, 4, 5, 10, 20, 71, 142, 284, 355, 710\}$$

and the minimal elements (as will always be the case) are just the prime divisors of 1420, namely 2, 5, and 71 and the maximal elements (as always will be the case) are just 1420 divided by the prime divisors of 1420, namely $1420/2 = 710$, $1420/5 = 284$, and $1420/71 = 20$. This is represented in Figure 2.5 since the vertices which are closest to 1, i.e. adjacent to 1, are indeed 2, 5, and 71; while the vertices closes to 1420 are indeed 710, 284, and 20.

Exercise 2.60. Suppose that $(X, \preccurlyeq)$ is a poset and that X is a finite, non-empty set. Prove that there exists maximal and minimal elements.
Hint! This should be obvious, just look at what it means for an element of X to not be maximal (or minimal).

Solution

Of course the above exercise is false if X isn't finite. For instance, if $X = \mathbb{Z}$, and we take $\preccurlyeq = \leq$, the usual less than or equal to relation, there is neither a maximal or minimal element (of course infinity is not a number!!!!).

Definition 2.76. Let $\sim$ be an equivalence relation on a set X. Let $a \in X$. We define

$$a/_\sim \stackrel{\text{def}}{=} \{b \in X : a \sim b\}$$

called *the equivalence class of a mod $\sim$* or simply the *equivalence class of a* or even *the class a*. We also define

$$X/_\sim \stackrel{\text{def}}{=} \{a/_\sim : a \in X\}$$

called X mod $\sim$ or the *quotient of X by $\sim$*.

Remark 2.77.

(1) By definition, an arbitrary element of ${}^{X}/_{\sim}$ is equal to ${}^{x}/_{\sim}$ for some $x \in X$. Therefore for proofs involving proving something about ${}^{X}/_{\sim}$ we typically start by saying

Let ${}^{x}/_{\sim}$ $(x \in X)$ *be an arbitrary element of* ${}^{X}/_{\sim}$.

(2) By definition equivalence classes ${}^{x}/_{\sim}$ are subsets of X. However it is often "better" to find other ways of thinking about these classes based on the context of the problem. On the other hand note that ${}^{X}/_{\sim}$ is not a subset of X (but it is a subset of the power set of X).

(3) The set ${}^{x}/_{\sim}$ is never empty since we always have that $x \in {}^{x}/_{\sim}$ because $x \sim x$ by reflexivity. Indeed, we have a natural surjective function[III] denoted by $\pi_{\sim}$

$$\begin{aligned} \pi_{\sim} &: X \to {}^{X}/_{\sim} \\ &: x \mapsto {}^{x}/_{\sim}. \end{aligned}$$

This is called the *quotient function (or map)*.

Exercise 2.61. Define a relation $+c$ on the set of all functions on $\mathbb{R}$, $\mathcal{F}_{\mathbb{R}}$, by $(f) \; +c \; (g)$ iff there exists a constant r real number so that $f + r = g$. Show that $+c$ is an equivalence relation. The class of f under the equivalence relation is ${}^{f}/_{+c} \stackrel{\text{def}}{=} f + c$.

SOLUTION

Example 2.78.

(1) If $\sim$ is the trivial equivalence relation on X which relates things only if they are are equal in the first place, i.e. $\sim = \mathbf{1}_X$, then for any $x \in X$ we have ${}^{x}/_{\sim} = \{x\}$ and $\pi_{\sim}$ is the bijection $x \mapsto \{x\}$. For this particular $\sim$, we always identify ${}^{X}/_{\sim}$ with X itself. Note this identification is the same identification from Convention 2.18 where we identify X with $\mathcal{P}_1(X)$.

(2) In the other extreme, if $\sim$ is the equivalence relation which relates any two elements of X, i.e. $\sim = X \times X$, then for any $x \in X$, we have ${}^{x}/_{\sim} = X$, ${}^{X}/_{\sim}$ consists of this single element, and $\pi_{\sim}$ is the constant function.

(3) (Now onto some more interesting examples.) If we again consider $\mathcal{Y}$ as the set of all people living in the US (ok, in a state in the US), and define $P_1 \sim P_2$ iff P_1 and P_2 live in the same state (which is clearly an equivalence relation), we have that

$${}^{\text{Mark Richard O'Brien}}/_{\sim} = \text{All people living in IL} = {}^{\text{Dick Durbin}}/_{\sim},$$

and we also have that ${}^{\mathcal{Y}}/_{\sim}$ can be identified with the set of all states.

(4) Let $\sim$ be the relation $a \sim b$ iff $a^2 = b^2$ defined on the set of all real numbers. This is an equivalence relation. For this relation we have that

$${}^{a}/_{\sim} = \{a, -a\}.$$

[III] Proving surjectivity is so obvious it hurts! Indeed, every element of ${}^{X}/_{\sim}$, by definition, is equal to ${}^{x}/_{\sim}$ for some $x \in X$ and we have that $\pi_{\sim}(x) = {}^{x}/_{\sim}$ by definition.

(5) Let $\asymp$ be the parallel relation from 1 back on page 50. Then for a non-zero vector **v**, the class of **v** can be interpreted as the line through the origin (missing **0**) which contains **v**.

(6) Suppose that $\sim$ is the equivalence relation in Figure 6 on the set $\{1, 2, 3, 4, 5, 6\}$.

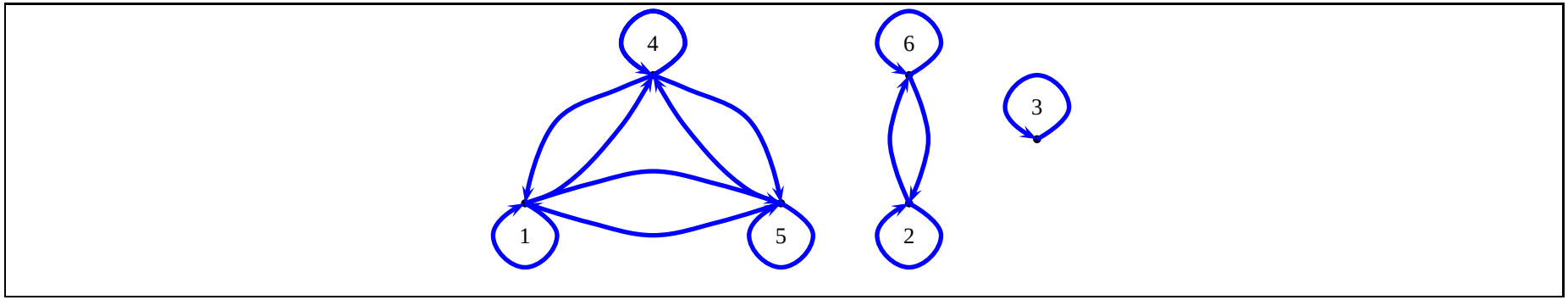


Figure 2.6: Figure 6

$$^1/_\sim = \{1, 4, 5\} \qquad ^2/_\sim = \{2, 6\} \qquad ^3/_\sim = \{3\}$$
$$^4/_\sim = \{1, 4, 5\} \qquad ^5/_\sim = \{1, 4, 5\} \qquad ^6/_\sim = \{2, 6\}$$

and

$$^{\{1,2,3,4,5,6\}}/_\sim = \{\{1, 4, 5\}, \{2, 6\}, \{3\}\}.$$

Moreover

$$\pi_\sim(1) = \{1, 4, 5\} \qquad \pi_\sim(2) = \{2, 6\} \qquad \pi_\sim(3) = \{3\} \qquad \text{etc. } (11)$$

In Example 2.72**(2)** on page 50 we showed that the relation $a \equiv_n b$ (which by definition holds iff $n \mid (a - b)$) was an equivalence relation.

DEFINITION 2.79. The equivalence class of a number a under this relation is denoted as $[\, a \,]_n$ (instead of $a/\equiv_n$). The set of equivalence classes are denoted as $\mathbb{Z}_n$ (instead of $\mathbb{Z}/_{\equiv_n}$). The set $\mathbb{Z}_n$ is known as the set of *integers mod n*.

As such, by the basic property of equivalence relations, we have

$$[\, a \,]_n = [\, b \,]_n \Leftrightarrow a \equiv_n b \Leftrightarrow n \mid (a - b).$$

In particular, we have that

$$[\, a \,]_n = [\, 0 \,]_n \Leftrightarrow n \mid a.$$

Example 2.80. We have that $\mathbb{Z}_2$ contains exactly two classes:

$$[\,0\,]_2 = \text{Set of all even integers} = \{\ldots, -6, -4, -2, 0, 2, 4, 6, \ldots\}$$
$$\text{and } [\,1\,]_2 = \text{Set of all odd Integers} = \{\ldots, -5, -3, -1, 1, 3, 5, 7, \ldots\}$$

For instance, see Exercise 2.47. This is because 2 divides the difference of two numbers iff both numbers are both even or both odd.

For other integers n you can think of $\mathbb{Z}_n$ as a generalization of cutting the integers up into two stacks: one stack containing the ven integers and the other stack containing the odd integers. It will turn that, in general, $\mathbb{Z}_n$ (for $n \neq 0$) contains exactly $|\,n\,|$ elements, [IV] with exactly one for every integer between 0 and $|\,n-1\,|$. We will probably talk so much about these $\mathbb{Z}_n$ dudes so much it will probably make your head spin, c.f. Section 3.4. One could safely argue that the this is the most important concept you learn in this course if you are not interested in going into group theory. Notice we have some quick facts:

(1) Since the integer 1 divides every number, we have that for any two integers a and b, $a \equiv_1 b$. As such for any integers a and b we have that

$$[\,a\,]_1 = [\,b\,]_1 .$$

In particular, any two integers are related and $\mathbb{Z}_1$ contains only 1 element.

(2) Every integer divides 0 but no number is divisible by 0. So

$$a \equiv_0 b \Leftrightarrow 0 \mid (a-b) \Leftrightarrow a = b.$$

As such,

$$[\,a\,]_0 = \{a\}$$

and $\mathbb{Z}_0$ can be identified with $\mathbb{Z}$ via

$$a \mapsto [\,a\,]_0$$

(3) Say an integer divides an integer is the same as saying that the negative of that integer divides it. Therefore we have that for all integers n and m

$$[\,a\,]_n = [\,a\,]_{-n}$$

As such $\mathbb{Z}_n = \mathbb{Z}_{-n}$.

In summary of these facts, studying $\mathbb{Z}_0$ is the same thing as studying $\mathbb{Z}$ and studying $\mathbb{Z}_n$ is the same as studying $\mathbb{Z}_{-n}$. As such it's reasonable to make the following convention:

Convention 2.81. For $\mathbb{Z}_n$ we always assume that n is a positive integer.

The following theorem states that the classes under any equivalence relation *partition* the set in the sense that every element of X is contained in exactly class.

Theorem 2.82. *Let $\sim$ be an equivalence relation on a set X. Then, the following are equivalent* (TFAE)

[IV] Granted that the elements of the set are themselves sets with an infinite number of elements, but this is not what I'm counting. For instance, for all intents and purposes, a person contains (damn well near) an infinite number of molecules, but you certainly wouldn't consider this if I asked you how many people are taking this course with you!

(1) $a \sim b$; [V]

(2) $a/\sim \bigcap b/\sim \neq \varnothing$;

(3) $a/\sim = b/\sim$;

(4) $a/\sim \subseteq b/\sim$;

(5) $b \in a/\sim$;

(6) $a \in b/\sim$.

REMARK 2.83. Therefore we have

$$a \sim b \Leftrightarrow a/\sim = b/\sim.$$

This is one of only several theorems that should be "internalized" in the sense that you should be able to jump between both sides if the iff at a moments notice. It is the mathematical way of explaining how somebody can call two different Nissan Maxima's the same car.

PROOF. We only prove the equivalence of the first three:**(1)**,**(2)**, and **(3)**. This is because **(5)** and **(6)**, are by definition, equivalent to the requirements that $a \sim b$ and $b \sim a$, respectively. We will pass by the fourth, **(4)** on the way.

1 ⇒ 2 Suppose that $a \sim b$. We wish to prove the equality of the two sets

$$a/\sim \quad \text{and} \quad b/\sim.$$

We do this in the usual way by showing that each is contained in the other. To this end let $x \in b/\sim$ be arbitrary. Then by definition of the class of b, this means that $b \sim x$. But since $a \sim b$ by the assumption **(1)** we can make use of the transitive property – which holds for $\sim$ since it is an equivalence relation – to conclude $a \sim x$. Therefore, by definition of the class of a, this means that $x \in a/\sim$. Since x was arbitrary we have shown that

$$b/\sim \subseteq a/\sim.$$

Really, what we have shown is

$$a \sim b \Rightarrow b/\sim \subseteq a/\sim. \tag{12}$$

But by the symmetric property we know that $b \sim a$ holds iff $a \sim b$. Therefore we can exchange a by b in Equation (12) to conclude the opposite inclusion

$$a/\sim \subseteq b/\sim.$$

Since mutual containment holds the two sets are equal as desired.

3 ⇒ 1 Suppose that we have the equality

$$a/\sim = b/\sim.$$

Then since $b \sim b$ by the reflexive property, we have that

$$b \in b/\sim = a/\sim \Rightarrow b \in a/\sim.$$

But by definition this means that $a \sim b$ as desired.

[V] By symmetry of the equivalence relation $\sim$, this is equivalent to $b \sim a$.

3 ⇒ 2 Since classes are never empty (the class of an element at the very least contains that element), the equality

$$[a]_\sim = [b]_\sim.$$

obviously implies the intersection is non-empty, i.e. **(2)**. For instance, we know that the element a is in both sets.

2 ⇒ 1 We are now assuming that

$$[a]_\sim \bigcap [b]_\sim \neq \varnothing.$$

This means

$$\exists x \in [a]_\sim \bigcap [b]_\sim$$

$$\begin{aligned}
&\exists x \text{ s.t. } x \in [a]_\sim \text{ and } x \in [b]_\sim \\
&\exists x \text{ s.t. } a \sim x \text{ and } b \sim x && \text{(Def'n of } [a]_\sim \text{ and } [b]_\sim) \\
&\exists x \text{ s.t. } a \sim x \text{ and } x \sim b && (\sim \text{ is symmetric}) \\
&a \sim b && (\sim \text{ is transitive})
\end{aligned}$$

But by the symmetric property, this means that $a \sim x$ and $x \sim b$. Lastly by the transitive property of the equivalence relation, we can stick these two together to conclude that $a \sim b$, which is **(1)**. ■

Therefore while the two elements a and $b \in X$ in a set are just related, in the new set $X/\sim$ the two classes are **equal**.

REMARK 2.84.

(1) Let $\sim$ denote the relation on $\mathbb{Z} \times \mathbb{Z}^\times$ from Exercise 2.49 back on page 51. We can use $\sim$ to give a definition of the rational numbers $\mathbb{Q}$ (from only the integers) by *defining*

$$[(a,b)]_\sim \overset{\text{def}}{=} \frac{a}{b} \quad \text{and} \quad \mathbb{Q} \overset{\text{def}}{=} \mathbb{Z} \times \mathbb{Z}^+ / \sim.$$

This is because we have that

$$(a,b) \sim (c,d) \Leftrightarrow \frac{a}{b} = \frac{c}{d}.$$

This is something that should feel rather familiar to you! For instance,

$$\frac{16}{32} = \frac{2}{4} = \frac{1}{2} \quad \text{and} \quad \frac{2,728,908}{482,664} = \frac{147}{26}.$$

In this context, (as will be the case quite often, but not all of the time) there is a natural representative for every class. Namely every rational number is equal to a unique rational number whose number and denominator are relatively prime. There are of course several reasons why we even bother to deal with fractions which are not reduced.

- It's much easier to simply write down a ratio of two large numbers such as

$$\frac{2,728,908}{482,664}$$

without worrying if the ratio is reduced or not[VI]. It seems as if students have been taking this to the extreme as of late. On every test I grade, I see students leaving answers like $\frac{4}{2}$ without reducing to 2. Even more extreme, I see things along the lines of $\frac{5}{1}$. While these nuances are annoying, I certainly cannot take off any points from the students work since what they wrote is certainly correct and the exams are not exams to test the students skills are reducing fractions.

- When performing operations with rational numbers we are naturally forced to leave the "reduced" world. For instance, in the calculation

$$\frac{4}{15}+\frac{7}{12}=\frac{16}{60}+\frac{35}{60}=\frac{51}{60}=\frac{17}{20}.$$

We changed the original fractions into two "compatible" fractions, added them by adding numerators, and then we had to reduce back down in the final step.

(2) Another situation where this shows up is in the addition of arrows (vectors). In most Calculus 3 textbooks, arrows are defined as an ordered pair of points in (say) the plane. The first point is to be pictured as the tail of the arrow, the other point was pictured as the head of the arrow. Two arrows were said to be equivalent if you can translate one on top of the other. This was the same as requiring that the two arrows have the same length (magnitude) and direction. In this context, the equivalence classes are called *vectors*. By translating the tail to the origin, every arrow is equivalent to some arrow based at the origin. These arrows play the same role as the reduced fractions from above. This is how we usually identify vectors with points – by identifying the arrow based at the origin with its endpoint. However, just like in $\mathbb{Q}$ it's often a good idea to leave this reduced world when performing operations. For instance, if we add two vectors, the resulting vector has tail at the origin and tip in the location where we would follow the first arrow and then *translate* the second arrow so it's tail is at the tip of the first arrow (hence this new arrow is not reduced!). Then the sum wound up having tip at the tip of the translated arrow.

The difference of two vectors was even stranger since the usual way of picturing it has tail at the tip of one of the arrows and tip at the tip of the other. In particular we don't even visualize the difference as a reduced vector.

(3) Occasionally one can even forget that we are dealing with classes completely and work things quite abstractly. By definition, a triangle is a subset of some Euclidean space. However, in a high school geometry class when one is trying to prove something about a triangle you are only given certain properties of the triangle. Also every student draws their own triangle on the piece of paper (hardly a plane) that is sitting in front of them. What you really ended up proving about a triangle you really ended up proving about an equivalence class of triangles where two triangles are equivalent if they are *congruent*. Therefore, that triangle in your notebook was a representative of an equivalence class of triangles while your friend sitting in front of you had another representative of the same equivalence class of triangles. This is how the mathematician makes precise the notation of "distinct but the same." In theory, a person in Chicago with a triangle with side lengths 2, 2, and 3 units has the potential of proving exactly the same things as somebody in Moscow with a triangle with same measures. One might even be tempted to say that the two students are dealing with "the same" triangle. However the teacher could not possibly ask a question such as "how far away is your triangle from Paris" since two congruent triangles can have different answers to this question.

(4) If we change the equivalence relation, we change what we can prove in the situation. Now picture the same two students, one in Chicago one in Moscow, from the previous part, but instead their teacher is asking them to prove things about a triangle whose angles are $45°, 45°, 90°$. For instance, the teacher might ask them to prove that the ratio of the length of the hypotenuse to the length of one of the other sides is 2. Since these triangles need not be congruent any longer the possible things one can hope to prove are different. The teacher can no longer ask what the parameter of the triangle is. However the two triangles are instead similar. Thus the students are studying to representatives of the same class of triangles under the equivalence relation which calls two triangle equivalent iff they are similar.

EXERCISE 2.62. Suppose that $\mathcal{R}$ is a reflexive and transitive relation on a set X.

SOLUTION

(1) Show that the relation $\sim$ defined as

$$a \sim b \iff a \mathcal{R} b \text{ and } b \mathcal{R} a$$

defines an equivalence relation.

(2) Show that

$$a' \sim a,\ a \mathcal{R}\, b,\ b \sim b' \Rightarrow a' \ \mathcal{R}\ b'.$$

(3) Explain why it makes sense to define a new relation $\preccurlyeq$ on $X/\sim$ by declaring

$$a/\sim \preccurlyeq b/\sim$$

whenever $a \mathcal{R}\, b$.

2.4.1 Partitions

Definition 2.85. Let X be a set. We say that $\mathcal{A} \subseteq \mathcal{P}(X)$ partitions X if every element of X is contained in exactly one element of $\mathcal{A}$.

In this case, we obviously we have that (if X is finite)

$$\#X = \sum_{A \in \mathcal{A}} \#A$$

(where $\#X$ denotes the number of elements in the set X). We also say that X is a disjoint union of the elements of $\mathcal{A}$ which is written

$$X = \bigsqcup_{A \in \mathcal{A}} A.$$

Remark 2.86. More generally, for two disjoint sets A and B we write

$$A \bigsqcup B$$

to mean the set $A \cup B$. This is called the *disjoint union* of A and B. This is even defined in more generality when we (somehow) identify A and B with $\widetilde{A}$ and $\widetilde{B}$ respectively with the added assumption that

$$\widetilde{A} \cap \widetilde{B} = \varnothing.$$

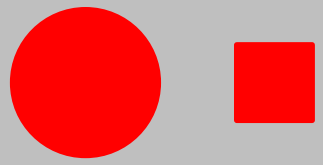
The most common way of doing this is by letting $\clubsuit$ be some object not in B and $\spadesuit$ be some object not in A and defining

$$\widetilde{A} := A \times \clubsuit \quad \text{and} \quad \widetilde{B} := \spadesuit \times B.^{\text{VII}}$$

Notice we have obvious identifications $A \to \widetilde{A}$ and $B \to \widetilde{B}$, but we now have that $\widetilde{A} \cap \widetilde{B} = \varnothing$. We define

$$A \bigsqcup B \overset{\text{def}}{=} \widetilde{A} \cup \widetilde{B}.$$

While this construction is a bit awkward (in my opinion), in practice it's easier than what has been said here. For instance, one can easily picture the disjoint union of the solid unit square and the solid unit circle (even though both sets contain the origin and as such are not really disjoint).



In practice, the most important property of a disjoint union is that the number of elements is additive, i.e.

$$\#\left(A \bigsqcup B\right) = \#A + \#B.$$

The set that is the United States can be partitioned in many different ways. It can be partitioned into voting districts, states, counties, regions, or even states beginning with a certain letter. However, it is not partitioned into cities since some places are not in any city. Partitions are really easy to come by. Just take a piece of paper and tear it to shreds!

We have already seen that equivalence relations partition sets into classes, but the converse is also true (in the inverse manner). This means that the concept of a partition is the same as the concept of an equivalence relation. More precisely,

THEOREM 2.87 (Fundamental Theorem of Mathematics[VIII])**.** *Let X be a set. Then*

$$\textit{Partions of } X \overset{1\text{-}1}{\leftrightsquigarrow} \textit{Equivalence Relations Defined on } X.$$

PROOF.

P's $\rightsquigarrow$ ER's Suppose that $\mathcal{A}$ is a partition of a set X. Define a relation $x \sim y$ iff x and y are in the same element of the partition. In more detail:

$$x \sim y \Leftrightarrow \exists A \in \mathcal{A} \text{ s.t. } x \text{ and } y \in A.$$

Or even better

$$x \sim y \Leftrightarrow \exists A \in \mathcal{A} \text{ s.t. } \{x, y\} \subseteq A.$$

I claim that this is an equivalence relation.

(a) (**Reflexive**) Let $x \in X$. By definition of a partition, $x \in A$ for some $A \in \mathcal{A}$. Thus $x \sim x$ and $\sim$ is reflexive.

(b) (**Symmetric**) Suppose that $x \sim y$. Then

$$\begin{aligned} x \sim y &\Rightarrow \exists A \in \mathcal{A} \text{ s.t. } x \text{ and } y \in A \\ &\Rightarrow \exists A \in \mathcal{A} \text{ s.t. } y \text{ and } x \in A \\ &\Rightarrow y \sim x. \end{aligned}$$

[VIII] This is something that the (yes **the**) algebraicist at Lycoming College, David Hayley, called it and it stuck with me!

(c) (**Transitive**) Suppose that $x \sim y$ and $y \sim z$. Then there exists $A_1 \in \mathcal{A}$ and $A_2 \in \mathcal{A}$ so that $\{x, y\} \subseteq A_1$ and $\{y, z\} \subseteq A_2$. In particular, $y \in A_1$ and $y \in A_2$. But, by definition of a partition, y is a **unique** element of the partition so $A_1 = A_2$. Hence $\{y, z\} \subseteq A_1$ which implies that x and z are both in A_1 and, as such, $x \sim z$ as desired.

ER's $\rightsquigarrow$ P's We've already seen how to get from equivalence relations to partitions: the equivalence classes of any equivalence relation defined on X partition X.

■

Exercise 2.63. Prove that the two processes defined in the above theorem, Theorem 2.87, are inverse to each other.
Hint! In one direction, you need to show that for any equivalence relation $\sim$ defined on X, that the relation

$$x \sim' y \Leftrightarrow \exists A \in X/\sim \text{ s.t. } \{x, y\} \subseteq A$$

gives us exactly the same relation $\sim$ that we started with. In the other direction, you need to show that for a partition $\mathcal{A}$ of X, the equivalence classes of the equivalence relation $\sim$ defined from $\mathcal{A}$ in the above proof has $X/\sim = \mathcal{A}$.

SOLUTION

Definition 2.88. Given a partition $\mathcal{A} \subseteq \mathcal{P}(X)$ we say that a subset $B \subseteq X$ is a set of representatives for $\mathcal{A}$ if

$$\begin{aligned} B &\to \mathcal{A} \\ b &\mapsto \text{The unique } A \in \mathcal{A} \text{ with } b \in A \end{aligned}$$

is a bijection.

Again, while the definition seems a bit yucky, the concept is really easy. We're simply choosing exactly one point for every set in the partition. For instance, if we partition the people in the US into the state they live in, then the set of governors form a set of representatives. Or perhaps we could take the set of oldest people in each state, or the set of tallest people or whatever.

2.4.2 Functions on Classes

As we have discussed extensively already, many things in mathematics are defined in terms of classes. However, in these situations one needs to be careful when defining functions on classes. For instance, it does not make any sense to evaluate an integral at point since a single function has many different integrals and there is not a natural one to choose in general. Another example is if we defined the relation $\mathcal{R}$ by

$$\frac{a}{b} \overset{\mathcal{R}}{\mapsto} \frac{a+1}{b}$$

also does not define a function on $\mathbb{Q}$ since we have for instance,

$$\left(\frac{1}{1}\right) \overset{\mathcal{R}}{\mapsto} \left(\frac{2}{1}\right) \quad \text{and} \quad \left(\frac{2}{2}\right) \overset{\mathcal{R}}{\mapsto} \left(\frac{3}{2}\right)$$

but of course $\frac{1}{1} = \frac{2}{2}$ but $\frac{2}{1} \neq \frac{3}{2}$.

The more general situation is as follows: we have a function $F : X \to Y$ and an equivalence relation $\sim$ defined on X. We would like to define a function

$$\widetilde{F} : X/_{\sim} \longrightarrow Y$$

by defining $\widetilde{F}\left({}^{a}/_{\sim}\right) = F(a)$. In general this is just a relation however.

Definition 2.89. In the situation above, we define the relation $\widetilde{F}$ by

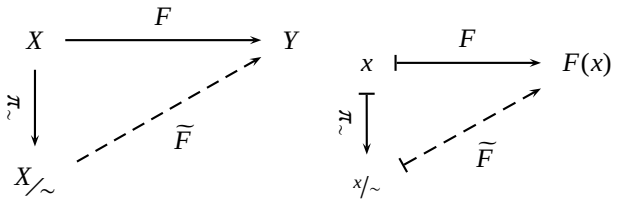
$$\{([x]_\sim, y) : F(x) = y\}.$$

This is called the *induced relation* of F. Notice as a relation it is the composition of the relations

$$\widetilde{F} = F \circ \pi_\sim^{-1}.$$

We will often denote this by

$$\widetilde{F}([x]_\sim) = F(x).^{\text{IX}} \tag{13}$$



Remark 2.90. We use the function notation from Definition 2.89 even though, in general, $\widetilde{F}$ is just a relation and not a function. This should be something you were "raised with." This leads to the complete nonsense statement that one often sees in many basic mathematics textbooks which says that

$$a = b \Rightarrow F(a) = F(b).$$

To see how nonsensical this statement is try this for a concrete set. For instance, for the natural numbers, we would have that $1 = 1$ but $F(1) \neq F(1)$ ... this is bs because $F(1)$ and $F(1)$ are identical (ok, perhaps down at the molecular scale the first $F(1)$ is different from the second one .. but ... yeah) so they must be equal! If we try to make this seem less trivial we start sounding a little nuts. For instance, if we said $1 = 3$ but $F(1) \neq F(3)$ we would be a lot of fun at parties (ok, the parties I go to 🙁) but we wouldn't make a lot of mathematical sense. This is a little less painful when defining induced relations on classes of an equivalence relation because we can conceive of things that equal but not identical, i.e $[a]_\sim$ can equal $[b]_\sim$ without having written the identical thing twice, i.e. even when $a \neq b$. For instance, mod 3, the statement $1 = 3$ is (at least somewhat) valid. Thus one could conceivably say that $1 = 3$ but $F(1) \neq F(3)$. This goes with the unwritten rule (that will send any logician into a manic attack mode) in mathematics that identical is stronger than equal.

Definition 2.91. Given a function $F : X \to Y$ and an equivalence relation $\sim$ we say that the relation $\widetilde{F}$ is *well-defined* if it is a function.

Example 2.92. Suppose that we already have a function

$$f : X/_\sim \longrightarrow Y.$$

[IX] This doesn't really make complete sense, see Remark 2.90 below.

Then define F to be the composition $f \circ \pi_\sim$. Then we have that $\widetilde{F} = f$. In this context we say that F is the *pullback* of the function f.

EXAMPLE 2.93. Except in trivial cases, it's very difficult to find everyday examples of equivalence relations and functions whose induced relation is well-defined. For instance, if we define an equivalence relation $\sim$ on the set $\mathcal{Y}$ of yankees (I mean American citizens ... ok, American citizens living in one of the 50 states, i.e. not in DC or Peurto Rico, or Guam ... etc) given by $P_1 \sim P_2$ if P_1 and P_2 live in the same county then none of the following functions induce well defined functions:

- Age $\rightarrow$ People living in the same county have different ages!
- Weight $\rightarrow$ People living in the same county have different weights!
- Birthday $\rightarrow$ People living in the same county have different birthdays!

The list goes on and on. There are even a couple of exceptions to the function which assigns people that live in the same country their state. For instance, Bristol country stretches between Massachusetts and Rhode Island. What we are really trying to find is

> What characteristics do any two people who live in the same county have in common?

The most obvious characteristics are the of course the county they live in, but in this case the induced map would be the identity! Another is the country they live in (country! not county ... why are these words so close? Somebody was trying to drive somebody crazy at some point), however this would be the constant function. Obviously we could start with a function defined on counties and then of course this would be well defined (as in the previous example). For instance, we could define the function which assigns every country its population. Then the pullback of this function is the function which assigns every person the population of the country they live in.

At least half of what is needed for $\widetilde{F}$ to be a function is always satisfied; namely it is always defined everywhere. This is because every element of $X/\sim$ has the form $x/\sim$ for some $x \in X$. But then since F is a function $F(x)$ is defined and $\widetilde{F}(x/\sim) = F(x)$ [X] which implies that $\widetilde{F}(x/\sim)$ is defined and hence $\widetilde{F}$ is defined everywhere. To show that $\widetilde{F}$ is single-valued requires more.

THEOREM 2.94. *$\widetilde{F}$ is well defined* $\Leftrightarrow$ $x \sim y \Rightarrow F(x) = F(y)$.

PROOF.

$\Leftarrow$ Suppose that $\widetilde{F}$ is well defined. Moreover suppose that $x \sim y$. We wish to show that $F(x) = F(y)$. By the equality of **(1)** and **(3)** in Theorem 2.82, we have that

$$x/\sim \; = y/\sim.$$

Therefore, if $\widetilde{F}$ is to be a function equals must go to equals, i.e.

$$\widetilde{F}\,(x/\sim) = \widetilde{F}\,(y/\sim)\,, \tag{14}$$

and we would have to have

$$F(x) = \widetilde{F}(x/\sim) = \widetilde{F}(y/\sim) = F(y).$$

The first and last equality are the definitions of $\widetilde{F}$ and the middle equality is Equation (14) from the previous paragraph.

$\Rightarrow$ On the other hand, suppose that the condition is met. We have already proven that $\widetilde{F}$ is defined everywhere, so by definition of a function we only need to prove that it is also single-valued. The proof of this is essentially arrived at by simply turning around the implications of the proof of necessity given in the first part. I do this out anyway because I'm a nice guy. Indeed, suppose that

$$x/\sim \; = y/\sim.$$

[X] Again, we should really be using the relation notation

$$(x/\sim, F(x)) \in \widetilde{F}$$

instead of the function notation $\widetilde{F}(x/\sim)$.

This time we need to show that Equation (14) holds. Since their class are equal, again the equality of **(1)** and **(3)** in Theorem 2.82 applies to show that $x \sim y$. By assumption means that $F(x) = F(y)$. But by definition of $\widetilde{F}$ given in (13) this means that

$$\widetilde{F}({}^{x}/_{\sim}) = \widetilde{F}({}^{y}/_{\sim})$$

and $\widetilde{F}$ is single-valued, and hence a function, as needed. ∎

Remark 2.95. Unlike the above proof, we will typically only make implicit use of Theorem 2.82

$$x \sim y \Leftrightarrow {}^{x}/_{\sim} = {}^{y}/_{\sim}.$$

That is we will use this equivalence without explicitly refererencing the theorem. This is done, not to be a jerk, but to get you guys to be able to make this important theorem implicit to you. You should almost jump out of your shoes whenever you see $x \sim y$ and yell with joy "that means that their classes are equal, yay!!!!"

Example 2.96. Let $F : X \to Y$ be a function and consider the relation $\sim_F$ considered earlier in Example 8 back on page 49. Then the induced relation $\widetilde{F}$ is well defined. This is almost by definition, indeed if $x_1 \sim_F x_2$, then by definition of $\sim_F$ we have that $F(x_1) = F(x_2)$. In fact the induced function is actually injective. To see this suppose we had

$$\widetilde{F}\left({}^{x_1}/_{\sim_F}\right) = \widetilde{F}\left({}^{x_2}/_{\sim_F}\right).$$

To show that $\widetilde{F}$ is injective we need to show that

$${}^{x_1}/_{\sim_F} = {}^{x_2}/_{\sim_F}.$$

(If they end up equal they must start off equal!)

$$\begin{aligned} F(x_1) &= \widetilde{F}\left({}^{x_1}/_{\sim_F}\right) && \text{(By def'n of } \widetilde{F}) \\ &= \widetilde{F}\left({}^{x_2}/_{\sim_F}\right) && \text{(By assumption)} \\ &= F(x_2). && \text{(By def'n of } \widetilde{F}) \end{aligned}$$

Therefore $F(x_1) = F(x_2)$ and as such $x_1 \sim_F x_2$ by definition of $\sim_F$. But this means that

$${}^{x_1}/_{\sim_F} = {}^{x_2}/_{\sim_F}.$$

Example 2.97.

(1) The relation $f : \mathbb{Z}_8 \to \mathbb{Z}_4$ given by $f([\,a\,]_8) = [\,a\,]_4$ is well defined. This is because if $a \equiv_8 b$ holds iff $a - b$ is divisible by 8. But this means that $a - b$ is divisible by 4 (by transitivity of the relation $n \mid m$). Therefore $a \equiv_4 b$. It's worth pointing out that in this situation $f = \widetilde{F}$ where F is the function $F : \mathbb{Z} \to \mathbb{Z}_4$ taking the integer n to $[\,n\,]_4$.

(2) The relation $f : \mathbb{Z}_5 \to \mathbb{Z}_8$ given by $f([\,a\,]_4) = [\,a\,]_8$ is not well defined. This is because

$$[\,0\,]_4 = [\,4\,]_4 \text{ but } [\,0\,]_8 \neq [\,4\,]_8.$$

Exercise 2.64. Consider the relation

$$\begin{aligned}\pi_m^n &: \mathbb{Z}_n \to \mathbb{Z}_m \\ &: [\,a\,]_n \mapsto [\,a\,]_m.\end{aligned}$$

(1) Find a function $\mathbb{Z} \to \mathbb{Z}_m$ which induces π_m^n,
Clarification! There is only one obvious one.

(2) Show that π_m^n is well defined iff $m \mid n$.

Solution

Exercise 2.65. Suppose that $m \mid n$ and $n = dm$. Define

$$\begin{aligned}\mathbf{M} &: \mathbb{Z}_m \rightsquigarrow \mathbb{Z}_n \\ &: [\,a\,]_m \mapsto [\,da\,]_n\end{aligned}$$

(1) Prove that **M** is well defined.

(2) Prove that **M** is injective.

Solution

Exercise 2.66. Suppose that $\sim'$ is an equivalence relation defined on Y and $\sim$ is an equivalence relation defined on X. Prove that

$$\widehat{F}\,(x/_\sim) \stackrel{\text{def}}{=} F(x)/_{\sim'}$$

is well-defined iff for all $x, x' \in X$,

$$x \sim x' \Rightarrow F(x) \sim' F(x').$$

Solution

Exercise 2.67. Show that the image of F is always the same as the image of $\widetilde{F}$ (even if $\widetilde{F}$ is not a function and is just a relation).

Solution

Exercise 2.68. Suppose that we have a function $F : X \to Y$ and an equivalence relation $\sim$. Suppose that $a \sim b \Leftrightarrow F(a) = F(b)$ holds. Prove that $\widetilde{F}$

(1) is well-defined

(2) injective

Solution

(3) defines a bijection $X/\sim \to f(X)$.

Definition 2.98. A relation $\sim$ defined on a set is *coarser* than another relation $\approx$ also defined on the same set as $\sim$ is if (as sets) $\approx \subseteq \sim$.

Note that $\sim$ is coarser than $\approx$ iff

$$x \approx x' \Rightarrow x \sim x'.$$

For example the relation which relates two people living in the same state is courser than the relation which relates two people living in the same town which is courser than a relation which relates two people living in the same house.

Exercise 2.69. For two relations $\sim$ and $\approx$ defined on a set X, prove that $\sim$ is coarser than $\approx$ iff for every $x \in X$

$$x/\approx \subseteq x/\sim.$$ [XI]

Solution

Remark 2.99. Building on the previous exercise, Exercise 2.69, note that Example 2.96 says that given a function $F\colon X \to Y$ and an equivalence relation $\sim$ the induced relation, $\widetilde{F}\colon X/\sim \to Y$, is well defined iff $\sim_F$ is coarser than $\sim$. For example, if $X := \{-1, 0, 1\}$, $Y := \{0, 1\}$, and $F\colon X \to Y$ is defined via $x \mapsto x^2$, then

$$\sim_F = \{(0,0), (1,1), (-1,-1), (-1,1), (1,-1)\}.$$

Since the only 'non-trivial' part of this equivalence relation says that -1 and 1 are related to each other, the only equivalence relation properly contained inside of this one is $\{(0,0), (1,1), (-1,-1)\}$, which is nothing other than the identity relation, $\mathbf{1}_X$. Therefore, there are only two possible induced maps of F: the one coming from $\sim_F$, which is the bijection $\{0\} \mapsto 0$, $\{-1, +1\} \mapsto 1$, and from $\mathbf{1}_X$ which induces the same map F.

Example 2.100. Note that roughly speaking, given a relation $\mathcal{R}\colon X \rightsquigarrow Y$ it's harder for $X/\sim \to Y$ to be a function. However, it's easier for $X \to Y/\sim$, to be a function. In fact, this is exactly what one does in a calculus course! Indeed: if we denote

$$C(\mathbb{R}) \stackrel{\text{def}}{=} \{f \in \mathcal{F}_{\mathbb{R}} : f \text{ is continuous}\}$$

as the set of continuous functions $\mathbb{R} \to \mathbb{R}$ (i.e. exactly the kinds of functions you study in Calculus I) then we have the relation

$$\begin{aligned} \int &: C(\mathbb{R}) \to C(\mathbb{R}) \\ &: f(x) \mapsto \int f(x)\,dx \end{aligned}$$

is not a function since any given continuous function has more than one anti-derivative. However, all differ by a constant. This means that

$$\widehat{\int} : C(\mathbb{R}) \to C(\mathbb{R})/_{+c}$$

is a function where the relation $+c$ is defined in Exercise 2.61 back on page 55. This is something you know quite well! For instance, you would write

$$\int x\,dx = \frac{x^2}{2} + c$$

which in calculus didn't have too precise of a meaning, but in our current language means that *the* anti-derivative of x is the set of all functions which differ from $x^2/2$ by a constant.

2.5 A Word About Cardinality*

The cardinality of a set is the mathematicians { put blah blah}

We don't put a lot of focus on cardinalities for a number of reasons including the fact that we will nearly always be interested in finite sets. Hopefully this is a topic that you spend considerably more time on in your analysis class. We give a quick run through here simply to demonstrate how it fits in with our current theory. It is also rather magical since it always to define what a 'number is.'

Rant 2.101. This was actually a tipping point for me in my decision to change to math (from physics) when I started graduate school. I had a philosophy professor who, quite frankly, got be a little bit cocky towards the sciences, and in particular, physics. One day in class he made the statement that "mathematicians do not even know what a number is." However having learning about cardinalities in a previous math class I was easily able to refute his very wrong claim. Thinking about this I began to think how much more difficult it is to perform the same feat when given a typical physics question.

Using equivalence relations we can define the cardinality of a set by defining what it means for two sets to have the same cardinality. This is something I have already hinted at a few times, c.f. Rant 2.15 back on page 11. For instance, if we wanted to tell an alien civilization what the symbol 7 meant, we could show them 7 of something (probably 7 aliens since that's one thing they have certainly seen before!) Let's first show why this is possible (see my comments after Definition 2.59 for motivation):

Theorem 2.102. *Define a relation, $\approx$, on the class of all sets via*

$$X \approx Y \Leftrightarrow \exists \text{ a bijection } X \to Y.$$

Then $\approx$ is an equivalence relation.

Proof.

Reflexive Let X be a set. Then, the identity function, $\mathbf{1}_X : X \to X$, is a bijection. In particular there exists a bijection $X \to X$ and so $X \approx X$ and $\approx$ is **Reflexive**.

Symmetric Suppose $X \approx Y$. Then there exists a bijection $\Phi : X \to Y$. But then Φ^{-1} is a bijection $Y \to X$. In particular there exists a bijection $Y \to X$ and so $Y \approx X$ and $\approx$ is **Symmetric**.

Transitive Suppose $X \approx Y$ and $Y \approx Z$. There there exist bijection $\mathfrak{G} : X \to Y$ and $\Psi : Y \to Z$. Then $\Psi \circ \mathfrak{G}$ is a bijection from $X \to Z$. In particular there exists a bijection $X \to Z$ and so $X \approx Z$ and $\approx$ is **Transitive**. ∎

Definition 2.103. Let X and Y be two sets. We say that X and Y *have the same cardinality* if $X \approx Y$. We denote this as $\#X = \#Y$.

We can use this to define the number 7 (or any number we'd like) as follows:

SubDefinition 2.104. The number 7 is the cardinality of the set

$$\{\spadesuit, \diamond, \clubsuit, \heartsuit, \blacksquare, \bigstar, T\}.$$

Of course the problem with this definition is that it's far from canonical. Of course a more canonical definition would be to "define" the natural number n as the cardinality of the set of integers between 1 and n, $\overline{n}$. The problem with this of course is that this definition is circular. Another possiblity is to define this inductively by defining the 'number zero' as the cardinality of the empty set. Then, assuming n via a set X (i.e. $n := \#X$) has been defined, we define $n + 1$ as the cardinality of the set

$$X \cup \{X\}.$$

Therefore we have that

$$0 \stackrel{\text{def}}{=} \#(\varnothing) \quad 1 \stackrel{\text{def}}{=} \#(\{\varnothing\}) \quad 2 \stackrel{\text{def}}{=} \#(\{\varnothing\}, \{\{\varnothing\}\}), \ldots$$

This makes sense since for any set X we have that $X \neq \{X\}$. Alas, this still isn't exactly elegant since how do we define induction without referencing natural numbers (this is possible, but will take us too far afield)?

In a typical analysis course one also proves the following:

Theorem 2.105.

(1) *Each of the sets $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ have the same cardinality.*

(2) *None of the sets from the previous part have the same cardinality as $\mathbb{R}$ has.*

(3) *Each of the sets $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, (and more generally) $\mathbb{R}^n$ have the same cardinality.*

(4) *None of the sets from the previous part have the same cardinality as $\mathcal{F}_{\mathbb{R}}$ has.*

In fact one can define a relation $\preccurlyeq$ on the class of all sets by defining $X \preccurlyeq Y$ iff there exists an injection $X \to Y$. This is clearly not a partial order on the class of all sets since any two sets of the same cardinality as each $\preccurlyeq$ each other. However there obviously are distinct sets of the same cardinality. However $\preccurlyeq$ does descend to a partial order on the class of sets mod $\approx$, i.e. if $X \preccurlyeq Y$ and $Y \preccurlyeq X$, then X and Y have the same cardinality.[I] In this way we have

$$0 \preccurlyeq 1 \preccurlyeq \cdots \preccurlyeq 6 \preccurlyeq 7 \cdots \preccurlyeq 5280 \preccurlyeq \cdots \preccurlyeq \#\mathbb{N} \preccurlyeq \#\mathbb{R} \preccurlyeq \mathcal{F}_{\mathbb{R}} \preccurlyeq \mathcal{F}_{\mathcal{F}_{\mathbb{R}}} \preccurlyeq \cdots.^{\text{II}}$$

[I] Proving this is not as obvious as you think, but it's less difficult than you may fear!

[II] The continuity hypothesis, which is an axiom of set theory, roughly says that we haven't missed anything.

Exercise 2.70. For any set X prove that X and the power set of X, $\mathcal{P}(X)$, have different cardinalities.
Careful! No! The statement $n < 2^n$ is not a valid proof.
Hint! This not hard, it's just a bit difficult to wrap your brain around. Here's a possible outline for you to follow:

Step 1 Assume there exists a bijection $X \to \mathcal{P}(X)$ and denote it by $x \mapsto A_x$ (where $A_x \subseteq X$).

Step 2 Define B as $\{x \in X : x \notin A_x\}$.

Step 3 Why is there a $y \in X$ with $A_y = B$?

Step 4 Now following a similar argument used in Russell's paradox prove that both $y \in B$ and $y \notin B$ (y is as in the previous part and B as in two parts up).

SOLUTION

3 Rings (i.e. crap that can be added and multiplied)

While we will not make a formal study of rings until the appendix, Appendix E, we will give an informal look at some examples. Roughly, a ring is a set where one can add, subtract, and multiply in a reasonable way. Moreover rings are required to have something which behaves like the number 0 and something (else) which behaves like the number 1. For instance the integers, reals, complexes and rationals are rings. However the set of cats named "Tom" in the world is not a ring. We do not consider the natural number $\mathbb{N}$ to be a ring since subtraction is not defined in general. While we **do** require subtraction be defined, we do **not** require that division be defined. For instance the integers, $\mathbb{Z}$, make up a perfectly good ring.

We won't give a formal definition here because we will only be dealing with a handful of examples of rings, and again I point the reader to Appendix E–more precisely Definition E.1–for the formal definitions and a more organized treatment of the definitions which appear in this chapter. For all intents and purposes, whenever I ask to prove something for all rings you can simply look back at this chapter and go one by one through the examples of this chapter. In addition to the rings just mentioned we will introduce the quaternions, $\mathbb{H}$, and the integers mod n, $\mathbb{Z}_n$, which are rings which you are no doubt less familiar with (the latter has already been introduced, but we haven't discussed their arithmetic yet at this point). In the ring $\mathbb{Z}_n$, the class $[\,1\,]_n$ plays the role of 1 and $[\,0\,]_n$ plays the role of 0.

Temporary Definition 3.1. A ring is one of the sets $\mathbb{H}$, $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$, or one of the sets $\mathbb{Z}_n$ with $n > 1$.

As mentioned in the introduction to the notes, for a ring R we denote the set of non-zero elements of R as $R^\times$. We say that elements r and s of a ring *commute* if $rs = sr$ and they anti-commute if $rs = -sr$.[I]

[I] We will see that in the quaternions many pairs of elements anti-commute. For instance we have $\mathbf{ij} = -\mathbf{ji}$.

3.1 The Integers

Hopefully the material in this section is a review from a foundations course. We take the basic algebraic properties of the integers, including associativity, commutativity, and distributivity as given. We also take the basic properties of order as given.

First we define the notion of what it means for an integer to divide another integer so it takes a more central position (previously it was defined as part of an example.)

Definition 3.2. Let a and b be integers. Suppose there exists an integer q so that $b = qa$. Then we say

(1) a divides b, (or changing the focus from a to b)

(2) b is a multiple of a.

In this case, we use the notation $a \mid b$.

Note that if $a \mid b$, then $|b| \geq |a|$. In particular, if a and b are positive (which will be our assumption most of the time) we have that $b \geq a$.[I] Note also that, ignoring complications coming from 0, another way of saying that $a \mid b$ is to say that b/a is an integer.

The integers are denoted by $\mathbb{Z}$. For a fixed integer n, it is trivial to see that the subset of multiples n can be written in any of the following equivalent ways:

$$\{0, -n, n, -2n, 2n, -3n, 3n, \ldots\} = \{m \in \mathbb{Z} : n \mid m\} = \{na : a \in \mathbb{Z}\}. \tag{1}$$

Definition 3.3. This set is denoted $n\mathbb{Z}$ is defined as the common set in (1) .

So an integer is in $n\mathbb{Z}$ iff that integer is divisible by n. For instance, $2\mathbb{Z}$ is simply the set of even integers.

Lemma 3.4. $m\mathbb{Z} \subseteq n\mathbb{Z} \Leftrightarrow n \mid m$.

Proof.

$\Rightarrow$ Suppose $m\mathbb{Z} \subseteq n\mathbb{Z}$. Then since $m \in m\mathbb{Z}$ we have that $m \in n\mathbb{Z}$. But since $n\mathbb{Z}$ is the set of multiples of n this implies that $n \mid m$.

$\Leftarrow$ Suppose that $n \mid m$. Then since divisibility is transitive, for any integer N, if N is a multiple of m, then N is also a multiple of n. But since $n\mathbb{Z}$ is the set of multiples of n and $m\mathbb{Z}$ is the set of multiples of m, the previous statement is just a restatement of what we wanted to prove, i.e. that $m\mathbb{Z} \subseteq n\mathbb{Z}$. ∎

For subsets S and T of $\mathbb{Z}$ define their sum, $S + T$, as the set of all possible sums of elements in S and T. That is

Definition 3.5. $S + T \stackrel{\text{def}}{=} \{s + t : s \in S, t \in T\}$.

An important special case of this definition is

[I] This is because the q appearing in the definition would have to be positive and $b - a = (q - 1)a$, which implies that unless $q = 1$, $b - a > 0$ as well.

SUBDEFINITION 3.6. For any two integers n and m

$$n\mathbb{Z} + m\mathbb{Z} \stackrel{\text{def}}{=} \{na + mb : a, b \in \mathbb{Z}\}.$$

That is $n\mathbb{Z} + m\mathbb{Z}$ consists of all $\mathbb{Z}$–linear combinations of n and m.

If T contains only a single integer a, i.e. if $T = \{a\}$, then we denote the set $S + \{a\}$ as just $S + a$.

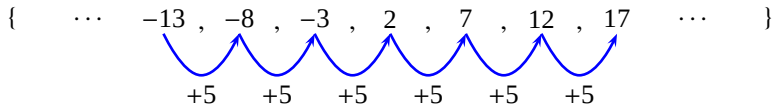
EXAMPLE 3.7.

(1) For any non-empty subset $S \subseteq \mathbb{Z}$ we have that

$$S + 0 = \{s + 0 : s \in S\} = \{s : s \in S\} = S.$$

(2) The set $2\mathbb{Z} + \{1\}$ is the set $\{2n + 1 : n \in \mathbb{Z}\}$, i.e. the set of odd integers. As with most sets consisting of a single element we will leave off the brackets around the singleton and simply denote this as $2\mathbb{Z} + 1$.

(3) The set $5\mathbb{Z} + 2$ is the set $\{5n + 2 : n \in \mathbb{Z}\}$ written increasing order is



In elementary school we would say that we are "counting by $5's$ starting from 2," but with the added catch that we also count this way in the negative direction as well.

(4) More generally, for integers m and a we have

SUBDEFINITION 3.8.

$$m\mathbb{Z} + a = \{mn + a : a \in \mathbb{Z}\} = \{a, a \pm m, a \pm 2m, a \pm 3m, \ldots\}. \tag{2}$$

Again, in elementary school we would say that we are "counting by $n's$ starting from a," but with the added catch that we also count this way in the negative direction as well. Of course, like for any set of integers, we have

$$n\mathbb{Z} + 0 = n\mathbb{Z}.$$

(5) If $S = \{0, 1, 2\}$ and $T = \{-1, 3, 5\}$, then

$$\begin{aligned} S + T &= \{0 - 1, 1 - 1, 2 - 1, 0 + 3, 1 + 3, 1 + 5, 2 + (-1), 2 + 3, 2 + 5\} \\ &= \{-1, 0, 1, 3, 4, 6, 1, 5, 7\} \\ &= \{-1, 0, 1, 3, 4, 5, 6, 7\}. \end{aligned}$$

Exercise 3.1. For subsets $S, T \subseteq \mathbb{Z}$ prove that $S + T$ is the image of the map

$$\begin{aligned} + : \mathbb{Z} \times \mathbb{Z} &\to \mathbb{Z} \\ : (a, b) &\mapsto a + b \end{aligned}$$

restricted to $S \times T$.
Hint! This is simply a definition check.

Exercise 3.2. Let S, T, and R be non-empty subsets of $\mathbb{Z}$. Prove

(1) $S + (T + R) = (S + T) + R$.

(2) $S + T = T + S$.

(3) $S + 0 = S = 0 + S$.

(4) $T \neq 0 \Rightarrow S + T \neq S$.

(5) $S + (T \cup R) = (S + T) \cup (S + R)$.

Now assume that S and T are finite. Prove

(6) $\#S \leq \#(S + T) \leq \#S \cdot \#T$

(7) Prove that the first inequality, $\#S \leq \#(S + T)$, from the previous part, Part **(6)**, is an equality iff $\#T = 1$.

(8) Consider the second equality, $\#(S + T) \leq \#S \cdot \#T$, from Part **(6)**

(a) Given an example where this inequality is

(i) Strict

(ii) Equality.

(b) Give a very rough sketch of what intuition you have developed of when the equality is strict or not.

Definition 3.9.

(1) A *greatest common divisor* (gcd) of integers a and b, denoted $\gcd(a, b)$ is an (unique up to sign) integer d which satisfies

gcd1 $d \mid a$ and $d \mid b$

gcd2 $\forall \widehat{d} \in \mathbb{Z}$, if $\widehat{d} \mid a$ and $\widehat{d} \mid b$ then $\widehat{d} \mid d$.

(2) A *least common multiple* (lcm) of integers a and b, denoted $\operatorname{lcm}(a, b)$ is an (unique up to sign) integer $\mathcal{L}$ which satisfies

lcm1 $a \mid \mathcal{L}$ and $b \mid \mathcal{L}$

lcm2 $\forall \widehat{\mathcal{L}} \in \mathbb{Z}$, if $a \mid \widehat{\mathcal{L}}$ and $b \mid \widehat{\mathcal{L}}$ then $\mathcal{L} \mid \widehat{\mathcal{L}}$.

Remark 3.10. We can generalize these definitions to the case of more than two integers in the following obvious way so, for instance, the greatest common divisor of a list of integers divides every number in the list and is divisible by any other number which divides every number in the list.

Exercise 3.3. Suppose a and b are integers. Prove that

(1) Show uniqueness of greatest common divisors up to sign: in other words show that if positive integers d_1 and d_1 satisfy **gcd1** and **gcd2**, then $d_1 = \pm d_2$. [II]

(2) Show uniqueness of least common multiples up to sign; in other words show that if positive integers $\mathcal{L}_1$ and $\mathcal{L}_2$ satisfy **lcm1** and **lcm2**, then $\mathcal{L}_1 = \pm\mathcal{L}_2$.

Hint! For both parts, you'll need to use the fact for integer m and n, if $m \mid n$ and $n \mid m$, then $n = \pm m$.

SOLUTION

This exercise means (after we prove existence) that any pair of numbers has a unique positive greatest common divisor. Really, we ought to right $\gcd(a, b) = \{d, -d\}$, but we will not do this. See Rant 3.14 just below for more on this. We make the following convention:

Convention 3.11. Unless we're clearly, we always take the gcd of two numbers to be positive.

Just try to keep in mind that when I say that $\gcd(a, b) = d$, technically I really mean that $\gcd(a, b) = d$ or $-d$.

Definition 3.12.

(1) Two integers, a and b, are *relatively prime* if the only positive integer which divides both a and b is 1. In this case we say that a is *prime to* b or a and b are *coprime*.

(2) A number other than 1 or -1 is *prime* if for every integer n, either

- $p \mid n$ or
- p and n are relatively prime.

(3) An integer which does not equal 1 or -1 is a *composite* if it is not a prime number.

Therefore numbers are characterized as either being prime, composite, or invertible (i.e. equal to $+1$ or -1).

In Theorem 3.31 we will show that greatest common divisors exist. Note that if $a \mid b$ (both positive), then a satisfies the necessary properties to be a greatest common divisor of a and b. Since greatest common divisors are unique by Exercise 3.3, this implies

Proposition 3.13. *If* $a \mid b$, *then* $\gcd(a, b) = a$. ∎

[II] Existence is proven by Theorem 3.31 below.

Therefore, the definition of a (positive) prime p requires for all integers m, that either $\gcd(p, m) = 1$ (if p does not divide m) or it equals p if it does.

EXERCISE 3.4. Prove that the only positive numbers which divide a (positive) prime number p are 1 and p.

SOLUTION

The presence of negative numbers can be quite annoying when studying divisibility. For instance, we have already seen that the greatest common divisor is technically not single valued due to these pesky negative numbers. For prime numbers, it should be clear from context whether or not we are allowing that integer to be negative or not. A common theme in our text is to ignore negative numbers and leave the annoying part to you.

RANT 3.14. In reality, we could define an equivalence relation where we relate any two numbers if their absolute values are the same. In which case nearly all annoyance disappears when we deal with classes instead of just the numbers themselves. But since you are probably still scared to death of equivalence classes (why?? They are your friends !! ⌢̈) we won't do this. Historically, this kind of procedure ended up leading to the notion of *ideals* which are one of the most important concepts in ring theory. In fact you can check that for all notions of divisibility we can replace a number n by an *ideal number* $\langle n \rangle$ which is really the pair of numbers $\{+n, -n\}$.

The basic axiom about the natural numbers we shall use is the *well ordering principle.*

THEOREM 3.15 (Well Ordering Principle)**.** *Every non-empty subset of* $\mathbb{N}$ *contains a least element.*

This is equivalent to the principle of induction.

THEOREM 3.16 (Induction)**.** *Suppose* $S \subseteq \mathbb{Z}^+$ *satisfies the following properties:*

- $1 \in S$*;*
- $n \in S \Rightarrow n + 1 \in S$*.*

Then $S = \mathbb{Z}^+$*.*

PROOF. Suppose that $S \neq \mathbb{Z}^+$. Then the set $\mathbb{Z}^+ \setminus S$ is non-empty and, by the well ordering principle, must contain a least natural number n. Since $1 \in S$ by assumption, we know that $n \neq 1$. Therefore $n - 1 \neq 0$. Since n is the *least* positve integer not in S, and $n - 1 < n$, we know that $n - 1 \in S$. But by the second assumption on the set S, we know that this implies that $n = (n - 1) + 1 \in S$. But this implies that n is both in S and not in S, a contradiction. ■

RANT 3.17. If I started doing chores and always had one more thing to do, I would never be done doing chores!

EXERCISE 3.5. Instead assume the principle of induction as an axiom and use it to prove the well ordering principle.

SOLUTION

As you are familiar induction is often used when S is a set of natural numbers where a certain statement is true. We will not review induction further since a foundations class is a prerequisite for this course.

Definition 3.18. A non-empty set $\mathcal{A} \subseteq Z$ is called a *subgroup* of $\mathbb{Z}$ if every $\mathbb{Z}$–linear combination of elements of $\mathcal{A}$ is in $\mathcal{A}$.

That is for all $a, b \in \mathcal{A}$ and for all $r, s \in \mathbb{Z}$ we have that

$$ra + sb \in \mathcal{A}.$$

Remark 3.19. Notice that if $\mathcal{A}$ is a subgroup of $\mathbb{Z}$, then since $\mathcal{A}$ is assumed non-empty, there exists an $n \in \mathcal{A}$. Therefore, since the number 0 can be written as $0 \cdot n + 0 \cdot n$, we know that $0 \in \mathcal{A}$ as well. Moving on, since we now know that $0 \in \mathcal{A}$ since every multiple of n can be written in the form

$$qn = q \cdot n + 1 \cdot 0$$

we know that every multiple of n is in $\mathcal{A}$ as well, i.e. $n\mathbb{Z} \subseteq \mathcal{A}$. In particular, since $-n = (-1)n$, we know that if $n \in \mathcal{A}$ so is its negative, $-n$. If $n \neq 0$ we know that either n or $-n$ is greater than 0. Therefore, every subgroup of $\mathbb{Z}$ which contains an number other than 0 contains a positive number.

The following exercise gives us a two step process to check whether or not a set of integers forms a subgroup of $\mathbb{Z}$. It will be generalized by the Two Step Subgroup Test, Defintion 6.7**(??)** coming up on page **??**

Exercise 3.6. Show that non-empty set of integer $\mathcal{A}$ is a subgroup of $\mathbb{Z}$ iff

(1) $\forall n \in \mathcal{A},\ -n \in \mathcal{A}$;

(2) $\forall n, m \in \mathcal{A},\ n + m \in \mathcal{A}$.

SOLUTION

Example 3.20. For any integer $d \in \mathbb{Z}$, $d\mathbb{Z}$ is a subgroup of $\mathbb{Z}$ since if m_1 and m_2 are divisible by d, then so is $m_1 + m_2$ and $-m_1$. Therefore, by the previous exercise, $d\mathbb{Z}$ is a subgroup of $\mathbb{Z}$.

Example 3.21. Suppose that $\mathcal{A}$ and $\mathcal{B}$ are subgroups of $\mathbb{Z}$. Then $\mathcal{A} + \mathcal{B}$ is also a subgroup of $\mathbb{Z}$. Indeed, if

$$n \in \mathcal{A} + \mathcal{B} \Rightarrow \exists a \in \mathcal{A}, \text{ and } b \in \mathcal{B} \text{ s.t. } n = a + b.$$

But since $\mathcal{A}$ and $\mathcal{B}$ are subgroups of $\mathbb{Z}$, both $-a \in \mathcal{A}$ and $-b \in \mathcal{B}$. Hence

$$-n = -(a + b) = (-a) + (-b) \in \mathcal{A} + \mathcal{B}.$$

If moreover $m \in \mathcal{A} + \mathcal{B}$ as well, then there exists integers $a' \in \mathcal{A}$ and $b' \in \mathcal{B}$ so that

$$m = a' + b'.$$

But then we have

$$n + m = a + b + a' + b' = (a + a') + (b + b') = A + B.$$

Since $\mathcal{A}$ is a subgroup of $\mathbb{Z}$, and since $A = a + a'$, we have that $A \in \mathcal{A}$. Similarly $B \in \mathcal{B}$. But this means that $n = A + B \in \mathcal{A} + \mathcal{B}$.

Example 3.22. For any sequence of integers $n_1, \ldots, n_m \in \mathbb{Z}$ the set

$$n_1\mathbb{Z} + \cdots + n_m\mathbb{Z}$$

of integer linear combinations of

$$n_1, \ldots, n_m$$

is a subgroup of $\mathbb{Z}$. Indeed, we can use inductive use of the previous two examples. We record two special cases as a proposition:

Proposition 3.23. *For any $n, m \in \mathbb{Z}$ we have that $n\mathbb{Z}$ and $n\mathbb{Z} + m\mathbb{Z}$ are subgroups of $\mathbb{Z}$.* ■

Lemma 3.24. *If $r \in n\mathbb{Z} + a$, then $r - n$ is as well.*

Proof. Let $r \in n\mathbb{Z} + a$. Then, by definition, there exists a $q \in \mathbb{Z}$ so that $r = qn + a$. Hence $r - n = (q-1)n + a = q'n + a$. Therefore $r - n \in n\mathbb{Z} + a$ as well. ■

The two most important theorems about integers in this course are

Theorem 3.25 (The Division Algorithm). *For any two integers n and m there exists unique integers q and $r < |n|$ so that $m = qn + r$.*[III]

Remark 3.26. The proof of this is based on the algorithm you would more than likely follow to find this. For instance for the integers 1003 and 21, to find the q you would keep taking multiplies of 21 until you are finally greater than 1003. The last integer before you are greater than 1003 is your q. The remainder, as every elementary school student knows, is less than 21 (if it weren't you should have used $q + 1$ instead. In general, the well ordering principle guarantees us that such an integer q exists.

Proof. (Proof of Theorem 3.25) We prove only existence and leave uniqueness to you in Exercise 3.7. We assume that n and m are positive and let you tinker with the remaining cases. Let r be the least positive integer in the set

$$T \stackrel{\text{def}}{=} (n\mathbb{Z} + m) = \{m, \pm n + m, \pm 2n + m, \pm 3n + m, \ldots\}.$$

Such an integer exists since the set in question, T, contains a positive integer (since

$$m = 0 \cdot n + m \in n\mathbb{Z} + m$$

and m assumed positive), and, as such, the well-ordering principle applies. I claim that $r < n$. We prove this by contradiction. Therefore we assume $r \geq n$. Set r' to be the difference, $r - n$. Note that

- $r' < r$ since $n > 0$,
- $0 \leq r'$ since $r \geq n$ by assumption,
- $r' \in n\mathbb{Z} + m$ by the lemma.

Hence $r' \in T$ and is smaller than r, which is impossible since we assumed that r was the smallest integer in T. Hence $0 \leq r < n$ and $r = q'n + m \Rightarrow m = qn + r$ (with $q = -q'$) as desired. ■

[III] $|n|$ denotes the usual absolute value of the integer n.

Remark 3.27. In elementary school we would say that 104 divided by 7 is 14 with a remainder of 6. This is because, at that point, in elementary we still had no idea what a fraction was. In abstract algebra, the situation is similar. We want to express as many things as possible without leaving the set that we are in, in this case the integers. There are many reasons for this including elegance and insight. So while the statements

$$\frac{104}{7} = 14 + \frac{6}{7} \qquad \text{or as a mixed number} \qquad 14\frac{6}{7}^{\text{IV}}$$

are true, they are not statements in the integers. Thus it is better to write this as

$$104 = 14 \cdot 7 + 6$$

which is indeed a statement made entirely in the integers about the integers.

Notation 3.28. For integer n and m The unique integers q and $r < |n|$ so that

$$m = qn + r$$

will be denoted $q(n, m)$ and $r(n, m)$ respectively. If n if fixed we denote these as $q_{(n)}(m)$ and $r_{(n)}(m)$.

Example 3.29.

(1) Let $n = 1034$ and $m = 17$. Then we have

$$1034 = 60 \cdot 17 + 14; \qquad 14 < 17.$$

Therefore $17 = r(17, 1034)$ and $60 = q(17, 1034)$.

(2) If $n \mid m$, then we have that $r(n, m) = 0$ and $q(n, m) = \frac{m}{n}$.

(3) If $0 < a < n$, then since

$$a = 0 \cdot n + a, \ m < n$$

we have that $q_{(n)}(a) = 0$ and $r_{(n)}(a) = a$. In particular, for arbitrary positive integers, n and m, since

$$r_{(n)}(m) = r(n, m) < m$$

we have that

$$r_{(n)}(r_{(n)}(m)) = r_{(n)}(m).$$

Therefore we can think of $r_{(n)}$ as a projection $\mathbb{Z} \to \{0, \ldots, n-1\}$.

Corollary 3.30. *Let a and $n \neq 0$ be intgers. There exists a unique integer r so that*

[IV] Mixed numbers suck and will NEVER be used again!! ... ick!!

- $0 \leq r < n$
- $r \in n\mathbb{Z} + a$.

Moreover, this integer r is the smallest positive integer in $n\mathbb{Z} + a$

Proof. Just follow the proof of the division algorithm. ∎

Exercise 3.7. Prove the uniqueness of the integers r and q in the division algorithm.

SOLUTION

Theorem 3.31 (Euclidean Algorithm Formula)**.** *For any two integers m and n, we have*

$$\gcd(m, n)\mathbb{Z} = n\mathbb{Z} + m\mathbb{Z}.$$ [V]

Remark 3.32. Of course this doesn't work for $\mathbb{N}$! For instance, $\gcd(10, 25) = 5$ but obviously you can not get a nickel from only dimes and quarters!

Using Exercise 3.8 below, let's first give a corollary of the theorem (assuming it to be true) before we set out to prove it. The following corollary says that a and b are relatively prime iff there are integers $n, m \in \mathbb{Z}$ with

$$1 = na + mb.$$

Corollary 3.33. *Two integers n and m are relatively prime iff 1 is an integer linear combination of n and m.*

Proof.

⇒ Suppose a and b are relatively prime. Then, by Exercise 3.8 $\gcd(a, b) = 1$. But then by the theorem we know that

$$\mathbb{Z} = 1\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z} \Rightarrow 1 \in a\mathbb{Z} + b\mathbb{Z} \Rightarrow \exists\, n, m \in \mathbb{Z} \text{ s.t. } 1 = an + bm.$$

⇐ Suppose that $1 = an + bm$. Then for $d > 0$, we have

$$d \mid a, d \mid b \Rightarrow d \mid (an + bm) \Rightarrow d \mid 1 \Rightarrow d = 1.$$

As such a and b are relatively prime.

∎

Before proving this we prove a couple of lemmas. The first one says that if there exists some integer d so that

$$d\mathbb{Z} = n\mathbb{Z} + m\mathbb{Z}$$

then this dude d must in fact be the least common multiple. Then we go on to show that such a d actually exists.

Lemma 3.34. *Let n, m and d be positive integers. Then*

$$n\mathbb{Z} + m\mathbb{Z} = d\mathbb{Z} \Rightarrow d = \gcd(n, m).$$

Proof. Suppose

$$n\mathbb{Z} + m\mathbb{Z} = d\mathbb{Z}. \tag{3}$$

We show that d satisfies **gcd1** and **gcd2**.

gcd1 We need to show that $d \mid n$ and $d \mid m$. To do this, we only use one inclusion of our assumption, (3),

$$n\mathbb{Z} + m\mathbb{Z} \subseteq d\mathbb{Z}.$$

Since

$$n\mathbb{Z} + m\mathbb{Z} = \{na + mb : a, b \in \mathbb{Z}\}$$

we have for any integers a and b, $an + bm \in d\mathbb{Z}$, i.e.

$$\forall a, b \in \mathbb{Z} \quad (d \mid (an + bn)).$$

Since this holds no matter which two integers a and b we take, it is true if we take $a = 1$ and $b = 0$. Doing this we see that $d \mid n$. If we do the opposite, i.e. take $a = 0$ and $b = 1$, we see that $d \mid m$. That is d satisfies **gcd1**.

gcd2 We need to show that if $d' \mid n$ and $d' \mid m$, then $d \mid d'$. To do this, we use the other inclusion from (3)

$$n\mathbb{Z} + m\mathbb{Z} \supseteq d\mathbb{Z}.$$

Suppose $d' \mid n$ and $d' \mid m$. We need to prove that $d \mid d'$. Since $d' \mid n$ and $d' \mid m$ there exists integers q_n and q_m so that

$$n = q_n d' \quad \text{and} \quad m = q_m d'. \tag{4}$$

Since $d = d \cdot 1$, $d \in d\mathbb{Z}$ and since $d\mathbb{Z} \subseteq n\mathbb{Z} + m\mathbb{Z}$ we have that $d \in n\mathbb{Z} + m\mathbb{Z}$. Therefore, there exist integers a and b with $d = na + mb$. Plugging in via (4) we have

$$\begin{aligned} d &= (q_n d')a + (q_m d')b \\ &= (q_n a + q_m b)d' \\ &= qd'. \qquad\qquad (\text{With } q := q_n a + q_m b \in \mathbb{Z}) \end{aligned}$$

Therefore $d' \mid d$ as desired. ∎

Lemma 3.35. *For every subgroup $\mathcal{A}$ of $\mathbb{Z}$ there exists an integer $d \geq 0$ so that $\mathcal{A} = d\mathbb{Z}$.*

Remark 3.36. In words, this lemma says that every subgroup of $\mathbb{Z}$ consists only of multiples of some integer. We will be using this lemma throughout the textbook.

Proof. We need to show that there exists an integer D so that

$$D\mathbb{Z} = \mathcal{A}. \tag{5}$$

Clearly if $\mathcal{A}$ only contains 0, i.e. $\mathcal{A} = \{0\}$, then $\mathcal{A} = 0\mathbb{Z}$ and we are done (with $D = 0$). Therefore we can assume that $\mathcal{A}$ contains an integer other than 0. As we argued in Remark 3.19, this implies that $\mathcal{A}$ contains a positive integer. By the Well Ordering Principle, we know that $\mathcal{A}$ contains a least positive integer d. I claim that we can take $D = d$ in (5).

Let $N \in \mathcal{A}$ be arbitrary. We need to show that N is a multiple of d. By the division algorithm there exists integers q and $0 < r < d$ so that $N = qd + r$. Now since $N, d \in \mathcal{A}$, and since $\mathcal{A}$ is assumed to be a subgroup of $\mathbb{Z}$, we have that $q \cdot d \in \mathcal{A}$ and $N - q \cdot d \in \mathcal{A}$ as well. But this last term, $N - q \cdot d$, is equal to r. Therefore $r \in \mathcal{A}$.

We have shown that

(1) $0 \leq r < d$;

(2) $r \in n\mathbb{Z} + m\mathbb{Z}$.

Since d is least positive integer in $\mathcal{A}$, the only possibility left for r to be is 0. But this means that

$$N = q \cdot d + r = qd,$$

i.e. $d \mid N$ as desired. ∎

Proof. (**Proof of Theorem 3.25 [Version 1]**) By the previous lemma, Lemma 3.35, applied to the subgroup $n\mathbb{Z} + m\mathbb{Z}$ (it's a subgroup by Example/Theorem 3.23), we know that there exists a $d \in \mathbb{Z}$ so that

$$n\mathbb{Z} + m\mathbb{Z} = d\mathbb{Z}.$$

By the lemma before that, Lemma 3.34, we know that $d = \gcd(n, m)$ and we are done. ∎

Corollary 3.37. *The greatest common divisor of any two integers a and b exists. Moreover, it is the least positive integer which can be written as an integer combination of a and b, i.e.*

$$\gcd(a, b) = \min\{na + mb : na + mb > 0\}.$$ ∎

Example 3.38.

$$45 \cdot 13 - 8 \cdot 73 = 1$$

implies that 13 and 73 are relatively prime. It also implies that each of the pairs: $(45, 8)$, $(45, 73)$, $(13, 8)$ are also each relatively prime to each other.

Exercise 3.8. Prove that integers a and b are relatively prime iff $\gcd(a, b) = 1$.

Solution

Exercise 3.9. Suppose that $a' \mid a$ and $b' \mid b$. Prove that $\gcd(a', b') \mid \gcd(a, b)$ and $\operatorname{lcm}(a', b') \mid \operatorname{lcm}(a, b)$.

Solution

Exercise 3.10. Use the previous exercise, Exercise 3.9, to prove that if a and b are relatively prime, then every divisor of a is relatively prime to every divisor of b.

Solution

Exercise 3.11. Let n and m be any two integers. Suppose that d is some integer and $d \mid n$ and $d \mid m$.

(1) Prove that

$$\gcd\left(\frac{n}{d}, \frac{m}{d}\right) = \frac{\gcd(n, m)}{d}.$$

(2) Using **(1)** prove that n/d and m/d are relatively prime iff $d = \gcd(n, m)$.

Solution

The following exercise gives us one proof that the least common multiple of two integers exists

EXERCISE 3.12. Following the same basic outline as used for gcd of two integers, prove that

$$n\mathbb{Z} \cap m\mathbb{Z} = \text{lcm}(n, m)\mathbb{Z}.$$

SOLUTION

We construct a more concrete proof of Theorem 3.25.

LEMMA 3.39. *Suppose $d > 0$ satisfies*

(1) *$d \mid n$, $d \mid m$ and*

(2) *$d \in n\mathbb{Z} + m\mathbb{Z}$*

then $d = \gcd(n, m)$.

PROOF. We show that $n\mathbb{Z} + m\mathbb{Z} = d\mathbb{Z}$ and apply Lemma 3.34 to conclude that $d = \gcd(n, m)$. The assumption that $d \mid n$ and $d \mid m$ implies that there exist integers q_n and q_m so that $n = q_n d$ and $m = g_m d$.

$\subseteq$ Let $D \in n\mathbb{Z} + m\mathbb{Z}$ be arbitrary. Then there exists $a, b \in \mathbb{Z}$ so that $D = na + mb$. Therefore

$$\begin{aligned} D &= na + mb \\ &= d(q_n a + q_m b) \\ &= dq \in d\mathbb{Z}. \qquad (q := (q_n a + q_m b)) \end{aligned}$$

Therefore $D \in d\mathbb{Z}$. Since D was arbitrary, $n\mathbb{Z} + m\mathbb{Z} \subseteq d\mathbb{Z}$.

$\supseteq$ Since $d \in n\mathbb{Z} + m\mathbb{Z}$ we know that there are integers a and b so that $d = na + mb$. Let $D \in d\mathbb{Z}$ be arbitrary. Then there exists an integer q so that $D = qd$. As such,

$$\begin{aligned} D &= qd \\ &= q(na + mb) \\ &= na_q + mb_q \in n\mathbb{Z} + m\mathbb{Z}. \qquad (a_q := aq \text{ and } b_q := bq) \end{aligned}$$

Since D was arbitrary, $d\mathbb{Z} \subseteq n\mathbb{Z} + m\mathbb{Z}$. ■

PROOF. (**Proof of Theorem 3.25 [Version 2]**) Given integers n and m we repeatedly apply the division algorithm:

$$\begin{array}{lll} \text{Input } (m, n) & m = q_1 n + r_1 & (\text{Step 1 Output, } q, r_1 := q, r(n, m)) \\ \text{Input } (n, r_1) & n = q_2 r_1 + r_2 & (\text{Step 2 Output, } q, r_2 := q, r(r_1, n)) \\ \text{Input } (r_1, r_2) & r_1 = q_3 r_2 + r_3 & (\text{Step 3 Output, } q, r_3 := q, r(r_2, r_1)) \\ \vdots & & \vdots \\ \text{Input } (r_{i-2}, r_{i-1}) & r_{i-2} = q_i r_{i-1} + r_i & (\text{Step } i \text{ Output}, q, r_i := q, r(r_{i-1}, r_{i-2})) \\ \vdots & & \vdots \end{array}$$

That is assuming that r_j has been defined for all $j < i$ (and setting $r_{-1} := m$ and $r_0 := n$) we use the division algorithm to the pair (r_{i-2}, r_{i-1}) to get integers q_i and $r_i < r_{i-1}$ which satisfy

$$r_{i-2} = q_i r_{i-1} + r_i.$$

Since

$$r_i < r_{i-1} < r_{i-2} < \cdots < r_1 < a < b$$

the r's form a strictly decreasing sequence of integers ≥ 0. Therefore, we must eventually have that $r_N = 0$ for some positive integer N. I claim that $d := r_{N-1}$ is the greatest common divisor of n and m, i.e. I claim that

$$d = \gcd(n, m).$$

Looking at things from the bottom up we know that since $r_N = 0$ and since

$$r_{N-2} = q_i \overset{r_{N-1}}{\overset{\shortparallel}{d}} + \cancel{r_n}$$

we know that $d \mid r_{N-2}$. Moreover since

$$r_{N-3} = q_i r_{N-2} + d$$

r_{N-3} is the sum of two numbers which are divisible by d, and so r_{N-3} is also divisible by d. Continuing inductively in this way we see that d divides each r_i, in particular $d \mid n$ and $d \mid m$.

On the other hand, looking at things from the top down, we see that each $r_i \in n\mathbb{Z} + m\mathbb{Z}$. In particular, d, which equals r_{N-1}, is also in $n\mathbb{Z} + m\mathbb{Z}$. But we've already seen that if $d > 0$ satisfies

(1) $d \mid n, d \mid m$ and

(2) $d \in n\mathbb{Z} + m\mathbb{Z}$

then $d = \gcd(n, m)$. ■

EXAMPLE 3.40. Let's take $n = 1071$ and $m = 3675$. Then we have

Apply DA to pair $(3675, 1071)$	$3675 = 3 \cdot 1071 + 462$	(Step $1 \rightsquigarrow r_1 = 462$)
Apply DA to pair $(1071, 462)$	$1071 = 2 \cdot 462 + 147$	(Step $2 \rightsquigarrow r_2 = 147$)
Apply DA to pair $(462, 147)$	$462 = 3 \cdot 147 + 21$	(Step $3 \rightsquigarrow r_3 = 21$)
Apply DA to pair $(147, 21)$	$147 = 7 \cdot 21 + 0$	(Step $4 \rightsquigarrow r_4 = 0$).

Therefore, the process ended after 4 steps and the greatest common divisor is the output from the 3rd step: 21:

$$\gcd(1071, 3675) = 21.$$

3.2 Complex Numbers

We start out giving a definition of the complex numbers you may not be familiar with. It is what I like to call a "thinking backwards" definition since it will end up leading us to what we are more familiar with when dealing with complex numbers: numbers of the form $a + bi$. Simply *defining* the complex numbers as numbers of that form is not an acceptable definition because we do not know that i exists yet.

Definition 3.41. We identity the complex numbers, denoted $\mathbb{C}$, with pairs of real numbers, i.e. the plane $\mathbb{R}^2$.

That is a complex number is an ordered pair (z_1, z_2) where both z_1 and z_2 are real numbers.

Definition 3.42. For complex numbers $\mathsf{z} = (z_1, z_2)$ and $\mathsf{w} = (w_1, w_2)$ (so by definition z_1, z_2, w_1, and w_2 are real numbers, i.e. elements of $\mathbb{R}$) we define two operations:

(1) We define the (complex) *sum* of the complex numbers z and w, denoted $\mathsf{z} \oplus \mathsf{w}$, in the same way (in a non-fancy way) that we define addition on $\mathbb{R}^2$, i.e. component-wise:

$$\mathsf{z} \oplus \mathsf{w} = (z_1, z_2) \oplus (w_1, w_2) \stackrel{\text{def}}{=} (z_1 + w_1, z_2 + w_2)$$

where $+$ denotes the usual addition of the real numbers indicated.

(2) We define the (complex) *product* of the complex numbers z and w, denoted $\mathsf{z} \odot \mathsf{w}$, via

$$\mathsf{z} \odot \mathsf{w} = (z_1, z_2) \odot (w_1, w_2) \stackrel{\text{def}}{=} (z_1 \cdot w_1 - z_2 \cdot w_2, z_1 \cdot w_2 + z_2 \cdot w_2)^{\text{I}}$$

where $\cdot$ is the usual multiplication of the real real numbers indicated and (again) $+$ denotes the usual addition of the real numbers indicated.

Convention 3.43. We will (always) drop the circle around these operations and write $\odot$ as $\cdot$ (or simply via concatenation[II]) and $\oplus$ as $+$.

Example 3.44. For complex numbers $(3, 2)$ and $(1, 4)$ their sum is

$$(3, 2) + (1, 4) = (3 + 1, 2 + 4) = (4, 6).$$

and their product is

$$(3, 2) \cdot (1, 4) = (3 \cdot 1 - 2 \cdot 4, 3 \cdot 4 + 2 \cdot 1) = (3 - 8, 12 + 2) = (-5, 14).$$

[I] You are probably familiar with how to multiply complex numbers. We will see that this multiplication is the usual multiplication "worked" backwards. We use this "fancy" definition mostly for the sake of definition this operation in terms of operations we are already familiar with.

[II] That is for complex numbers we use either of the notations $\mathsf{z} \cdot \mathsf{w}$ or zw instead of $\mathsf{z} \odot \mathsf{w}$.

Example 3.45. Complex numbers of the form $(r, 0)$ behave just like real numbers in the sense that

- "Addition is addition," i.e.
$$(r, 0) + (s, 0) = (r + s, 0).$$
- "Multiplication of two of them is real multiplication," i.e.
$$(r, 0) \cdot (s, 0) = (r \cdot s - 0 \cdot 0, r \cdot 0 + 0 \cdot s) = (r \cdot s - 0, 0 + 0) = (r \cdot s, 0).$$
- More generally, "The product of $(r, 0)$ with a general $\mathsf{z} = (z_1, z_2) \in \mathbb{C}$is scalar multiplication," i.e.
$$(r, 0) \cdot (z_1, z_2) = (r \cdot z_1 - 0 \cdot z_2, r \cdot z_2 + 0 \cdot z_1) = (r \cdot z_1, r \cdot z_2) = r \cdot (z_1, z_2) = r \cdot \mathsf{z}.^{\text{III}}$$

Example 3.46. For the complex number $(0, 1)$ we have
$$(0, 1) \cdot (0, 1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0).$$

Motivated by the above examples, we introduce new notation:

Notation 3.47. For a real numbers r we write $(r, 0)$ as r. We write the complex number $(0, 1)$ as i.

Using the examples above we have

- For real numbers r and s no confusion can come about in whether the notation $r \cdot s$ means complex multiplication or (real) multiplication.
- For a real number r, the complex number $r \cdot i$ (as an ordered pair) is $(0, r)$.
- With our notation: $i^2 = -1$.

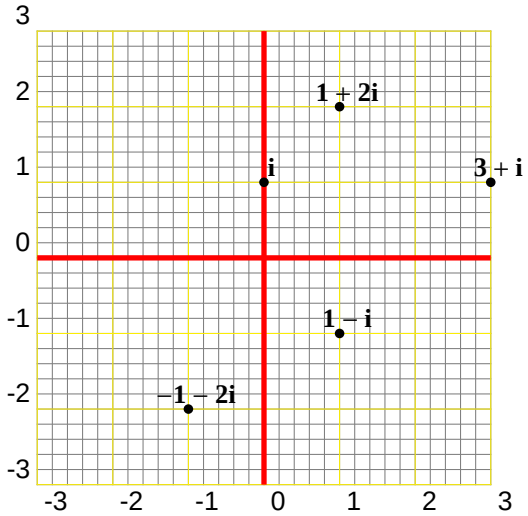
Using our (favorite!) new convention, we recover the usual way of writing a complex number
$$(a, b) = (a, 0) + (0, b) = a \cdot 1 + b \cdot i = a + b \cdot i.$$

Whenever we say that $\mathsf{z} = a + b \cdot i$ is a complex number we are implying that a and b are real numbers. We call a the *real part* of z and b the *imaginary part* of z.

Translating the definition of complex multiplication, 3.42, we also have
$$(z_1 + z_2 i) \cdot (w_1 + w_2\, i) = (\mathsf{z}_1, \mathsf{z}_2) \odot (\mathsf{w}_1, \mathsf{w}_2) \overset{\text{def}}{=} (z_1 \cdot w_1 - z_2 \cdot w_2, z_1 \cdot w_2 + z_2 \cdot w_2) = (z_1 w_1 - z_2 w_2) + (z_1 w_2 + z_2 w_2)\, i.$$

One advantage of thinking of complex numbers as points in the plane is that we now can "draw" subsets of the complex plane. For instance the point $1 + 2i$ would be put into the plane in the same spot that $(1, 2)$ would be put into the plane. For example, the points $1 + 2i$, i, $1 - i$, $-1 - 2i$, and $3 + i$ are indicated:

[III] The last $\cdot$ is the scalar multiplication which is defined in general for $r \in \mathbb{R}$ and $(v_1, \ldots, v_n) \in \mathbb{R}^n$ as
$$r \cdot (v_1, \ldots, v_n) = (rv_1, \ldots, rv_n).$$



We make this superficial distinction more explicit:

NOTATION 3.48. The set of complex numbers, denoted as $\mathbb{C}$, is the set of numbers

$$a + b\,i$$

where a and b are real numbers.

WARNING! 3.49. While this definition is what you should expect, it does mean that something along the lines of

$$\frac{1}{2+i}$$

is technically *not* a complex number. However, you can make it into one, namely

$$\frac{2}{5} - \frac{i}{5}.$$

We will tend to stick to our guns with that annoyance. One annoyance that we can not bare is that according to our definition

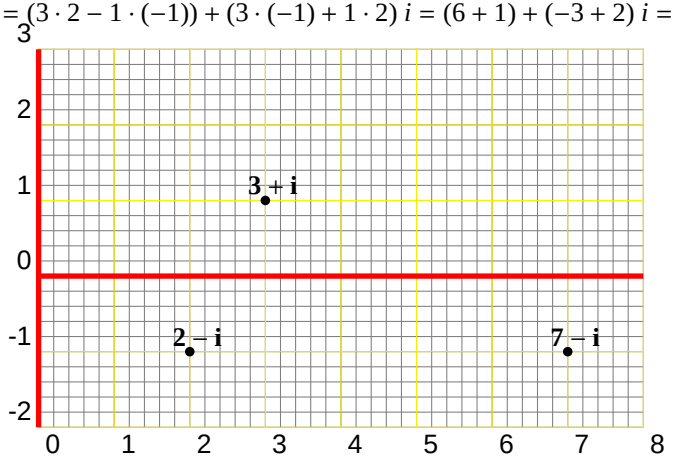
$$\frac{2+i}{3}$$

is also, technically, not a complex number (since it doesn't have the said form). However, we will be "sloppy" and not make a distinction between it and what we really ought to write it as:

$$\frac{2}{3} + \frac{i}{3}.$$

Example 3.50. Let $z = 3 + i$ and $w = 2 - i$. Then

$$z \cdot w = (3 \cdot 2 - 1 \cdot (-1)) + (3 \cdot (-1) + 1 \cdot 2)\, i = (6 + 1) + (-3 + 2)\, i = 7 - i.$$



Exercise 3.13. (Comparing Real and Imaginary Parts) Using only the definitions given show that for two complex numbers

$$z_1 + z_2 i = w_1 + w_2 i$$

holds iff $z_1 = w_1$ and $z_2 = w_2$.

Hint! If you read the definition correctly this is 100% trivial. There is no calculation that needs to be done! Simply use Definition 2.8 back on page 8.

Solution

Also note that we also have the following identities hold:

Theorem 3.51. *For complex numbers* z, z', z'', w, w', *and* w'' *the following hold:*

(1) (Distributive Properties)

$$z \cdot (w' + w'') = z \cdot w' + z \cdot w''$$
$$(z' + z'') \cdot w = z' \cdot w + z'' \cdot w.$$

(2) (Associative Properties)

$$(z \cdot z') \cdot z'' = z \cdot (z' \cdot z'')$$
$$(z + z') + z'' = z + (z' + z'').$$

(3) (Commutative Properties)

$$z + w = w + z$$
$$z \cdot w = w \cdot z.$$

Proof. We will prove a more general version of the distributive property when talking about quaternions. The properties about addition hold in any dimension. We will prove that multiplication is associative and leave it to you to prove that it is commutative. Write the complex numbers as

$$\mathsf{z} = z_1 + z_2\, i \qquad \mathsf{z}' = z_1' + z_2'\, i \qquad \mathsf{z}'' = z_1'' + z_2''\, i$$

Then,

$$\begin{aligned}(\mathsf{z}\cdot\mathsf{z}')\cdot\mathsf{z}'' &= ((z_1 + z_2\, i)\cdot(z_1' + z_2'\, i))\cdot(z_1'' + z_2''\, i)\\ &= \Big((z_1z_1' - z_2z_2') + (z_1z_2' + z_2z_1')\, i\Big)\cdot(z_1'' + z_2''\, i)\\ &= \Big((z_1z_1' - z_2z_2')z_1'' - (z_1z_2' + z_2z_1')z_2''\Big) + \Big((z_1z_1' - z_2z_2')z_2'' + (z_1z_2' + z_2z_1')z_1''\Big)\, i.\end{aligned}$$

Therefore the real part of this product is

$$z_1z_1'z_1'' - z_1z_2'z_2'' - z_1'z_2''z_2 - z_1''z_2z_2' \tag{6}$$

and the imaginary part of this (triple) product is

$$-z_2z_2'z_2'' + z_1z_1'z_2'' + z_1''z_1z_2' + z_1'z_1''z_2. \tag{7}$$

Similarly, if you expand the product $(\mathsf{z}\cdot\mathsf{z}')\cdot\mathsf{z}''$ (in the indicated order) you will see that you get the same result. (This should be apparent due to the highly symmetric nature of the equations, 6 and 7.) ■

Using the properties of addition and simply using the facts that

C1 $i^2 = -1$,[IV]

C2 for all real numbers a, $i\cdot a = a\cdot i$,

C3 and the distributive properties

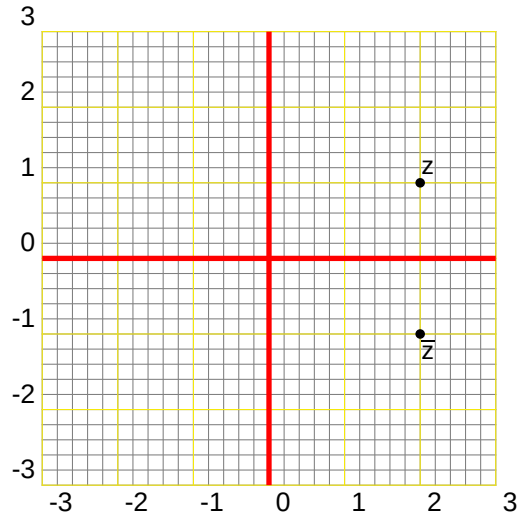
is enough to derive the multiplication used to define the complex numbers. Indeed using **only** these three properties we can completely recover

$$\begin{aligned}(a + b\, i)\cdot(c + d\, i) &= a\cdot(c + d\, i) + b\, i\cdot(c + d\, i)\\ &= ac + ad\, i + bc\, i + (b\, i)(d\, i)\\ &= ac + adi + bci + bd\, i^2\\ &= ac + adi + bci - bd\\ &= ac - bd + (ad + bc)\, i.\end{aligned}$$

Definition 3.52. For a complex number $\mathsf{z} = a + b\, i$ we define the *complex conjugate* of z to be $\overline{\mathsf{z}} := a - b\, i$.

In the complex plane, complex conjugation reflects the point z over the x–axis.

[IV] Powers of complex numbers have their obvious meaning, i.e. repeated multiplication.



We also have that for a complex number $\mathsf{z} = a + b\,i \in \mathbb{C}$

$$\begin{aligned}\overline{\mathsf{z}} \cdot \mathsf{z} &= (a - b\,i) \cdot (a + b\,i) \\ &= a^2 + b^2 + \cancelto{0}{(ab - ab)\,i} \\ &= a^2 + b^2 \\ &= \|\mathsf{z}\|^2\end{aligned}$$

where $\|\mathsf{z}\|$ is the length of $\mathsf{z} = a + b\,i$ when thinking of it as the vector (a, b) in the plane, $\mathbb{R}^2$. In particular this product is a non-negative real number which is equal to 0 iff $\mathsf{z} = 0$. This means for all non-zero complex numbers, we have that

$$\frac{\overline{\mathsf{z}}}{\|\mathsf{z}\|^2} = \mathsf{z}^{-1}$$

in the sense that it times z is equal to 1. For the usual unit circle, denoted as $\mathbb{S}^1$, is the subset of the plane,

$$\mathbb{S}^1 = \{(a, b) \in \mathbb{R}^2 : a^2 + b^2 = 1\} = \{(\cos\theta, \sin\theta) \in \mathbb{R}^2 : \theta \in \mathbb{R}\}.$$

It can be identified, via our identification of $\mathbb{R}^2$ with $\mathbb{C}$,[V] with the set of complex numbers

$$\{\mathsf{z} \in \mathbb{C} : \|\mathsf{z}\| = 1\} = \{\cos\theta + \sin\theta\,i \in \mathbb{C} : \theta \in \mathbb{R}\}.$$

[V] Again, this is

$$a + b\,i \in \mathbb{C} \overset{1\text{-}1}{\leftrightsquigarrow} (a, b) \in \mathbb{R}^2$$

3.2.1 Euler Formula

Perhaps the most awesome kick ass [VI] formula in the entire universe is the Euler formula. I found a special case of this formula,

$$e^{\pi i} = -1$$

while randomly typing in my calculator during my freshman astronomy class. After seeing it I rushed to the restroom and scrolled it on the bathroom wall. This is the one and only time that I have ever written on a bathroom stall in my life! A year or so later another student in a class with me happened to notice this written on the wall and upon returning from the restroom immediately asked the professor in an excited tone (in the middle of a lecture on planetary motion) about this very same formula. The astronomy professor was just as excited to tell us about this formula ... before we knew it we had a lecture about complex variables instead of Neptune. If we write the equation as $e^{\pi i} + 1 = 0$ we have the most important numbers, e, π, i, 1 and 0 in all of mathematics appearing in one equation! (I know what you're thinking, but you can't make a ringtone out of a math equation ....).

The general Euler Formula states that

$$e^{i\theta} = \cos\theta + \sin\theta\, i. \qquad \text{(Euler Formula)}$$

(This reduces to the formula already stated when we take $\theta = \pi$). [VII] In particular, we can express $\mathbb{S}^1$ as

$$\mathbb{S}^1 = \{e^{i\theta} : \theta \in \mathbb{R}\}.$$

(We only need to take θ to be between $0 \leq \theta < 2\pi$).

Exercise 3.14. Look up the history of this formula on Wiki at

http://en.wikipedia.org/wiki/Euler's formula.

SOLUTION

At the very least this formula contains every trig identity "on a stick." For instance

$$\begin{aligned} e^{ix}e^{iy} &= e^{i(x+y)} && \text{(See footnote)} \\ &= \cos(x+y) + \sin(x+y)\, i && \text{(Plug } x+y \text{ into (Euler Formula) for } x) \end{aligned}$$

while at the same time, simply using standard complex multiplication

$$(\cos x + \sin x\, i)(\cos y + \sin y\, i) = (\cos x \cos y - \sin x \sin y) + (\sin x \cos y + \cos x \sin y)\ i.$$

Thus if we compare real and imaginary parts (so we're using Exercise 3.13) of these two equations, we recover the usual double angle formulas

$$\cos(x+y) = \cos x \cos y - \sin x \sin y \qquad \text{and} \qquad \sin(x+y) = \sin x \cos y + \cos x \sin y$$

which are quite difficult to prove with a direct geometric proof. The polar coordinates of a point p in the plane is (r, θ) where r is distance of p to the origin and θ is the angle between the positive x-axis and the point. If (x, y) are the cartesian coordinates for p, then $x = r\cos\theta$ and $y = r\sin\theta$. Transferring this over to complex numbers this means that the corresponding complex point is

$$r\cos\theta + r\sin\theta\, i = r(\cos\theta + \sin\theta\, i) = re^{i\theta}.$$

[VI] Is kick ass hyphenated? kick-ass?

[VII] While the equation is fantabulously amazing, the most common proof kinda sucks: just equate the power series (the same power series for these functions one learns in Calculus 2) for both sides. When we stop to think about it, it makes sense that we would have to do this because what would it mean to take a number to an imaginary power if we didn't use a Taylor expansion? (What could i^i possibly mean?) The upshot of this is that most of the usual properties carry over to complex case, e.g. $e^{x+y} = e^x e^y$ holds for any two complex numbers x and y

Conversely if z is any point in $\mathbb{C}$, then $\|z\| = r$ and $\mathsf{z}/r \in \mathbb{S}^1$ which implies that there is some angle θ so that

$$\mathsf{z} = re^{i\theta}.$$

Moreover, provided that $z \neq 0$ and $0 \leq \theta < 2\pi$ this description is unique. This gives us a geometric description of multiplication by $e^{i\psi}$ for some angle ψ.

THEOREM 3.53. *If z is a complex number, then $e^{i\theta}\mathsf{z}$ is the point one gets my rotating z by an angle θ in the counter-clockwise direction.*

PROOF. We write

$$\mathsf{z} = r\,e^{i\theta'}$$

where r and θ' are the polar coordinates of z. Then

$$e^{i\theta}\mathsf{z} = r\,e^{i(\theta'+\theta)}.$$

Hence $e^{i\theta}\mathsf{z}$ is the same distance from the origin, but is now at an angle $\theta' + \theta$, i.e. it has been rotated by θ degrees. ■

EXAMPLE 3.54. Consider the point

$$\mathsf{z} = \sqrt{3} + i.$$

This is a point a distance

$$\|\mathsf{z}\| = \sqrt{\left(\sqrt{3}\right)^2 + 1^2} = \sqrt{4} = 2$$

from the origin and at an angle

$$\tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = \frac{\pi}{6}.$$

Therefore if we take z and multiply it by the complex number

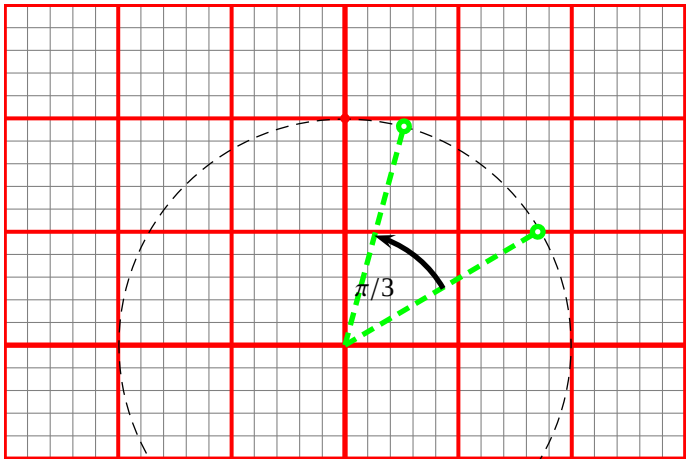
$$e^{\frac{\pi}{4}i} = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\,i$$

we'll end up with a point which is still a distance 2 from the orgin, but it now at an angle

$$\frac{\pi}{6} + \frac{\pi}{4} = \frac{5\pi}{12}.$$

To "see" this, let's draw it all out noting that a direct complex multiplication shows

$$\frac{\sqrt{2}}{2}\left(1+i\right)\cdot\left(\sqrt{3}+i\right) = \frac{\sqrt{2}}{2}\left[\left(\sqrt{3}-1\right) + \left(\sqrt{3}+1\right)\,i\right].$$



Another interesting example is titled DeMoivre's theorem,

THEOREM 3.55 (DeMoivre's Theorem)**.**

$$\cos nx + \sin nx\, i = (\cos x + \sin x\, i)^n .$$

PROOF. By Euler's formula we have

$$e^{inx} = \cos nx + \sin nx\, i \tag{8}$$

and, by the basic properties of the exponential function, we also have that

$$e^{inx} = \left(e^{ix}\right)^n . \tag{9}$$

Plugging in the Euler relation into this last equation, (9), and equating it to (8) gives the stated formula. ■

Again, the idea of the above derivation is trivial, we simply use $e^{nix} = (e^{ix})^n$ as an intermediary. The power comes after equating real and imaginary parts (so we're using Exercise 3.13 yet again) of this equation. In the $n = 2$ case we can use it to derive the familiar formula for $\cos 2x$ and $\sin 2x$ since

$$\begin{aligned}
\cos 2x + i \sin 2x &= (\cos x + \sin x\, i)^2 && \text{(DeMoivre w/ } n = 2) \\
&= (\cos x + \sin x\, i)(\cos x + \sin x\, i) \\
&= \left(\cos^2 x + 2\cos x \sin x + i^2 \sin^2 x\right) i && \text{(Expanding)} \\
&= \left(\cos^2 x - \sin^2 x\right) + \left(2 \sin x \cos x\right) i \\
&\Rightarrow \cos 2x = \cos^2 x - \sin^2 x \text{ and } \sin 2x = 2\cos x \sin x. && \text{(Exercise 3.13)}
\end{aligned}$$

Note, after applying the Pythogorean theorem to get rid of any $\sin^2$ terms, we could rewrite the formula for $\cos 2x$ in terms of just cosines:

$$\cos 2x = 2\cos^2 x - 1.$$

Note that this can be iterated:

$$\begin{aligned}\cos 4x &= \cos 2 \cdot 2x \\ &= 2\cos^2 2x - 1 \\ &= 2(2\cos^2 x - 1)^2 - 1 \\ &= 2(4\cos^4 x - 4\cos^2 x + 1) - 1 \\ &= 8\cos^4 x - 8\cos^2 x + 1.\end{aligned}$$

In general, if we define, $f(t) = 2t^2 - 1$, then $\cos 2^n x = f^n(\cos x)$ where f^n means f composed with itself n times (and **not** the n^{th} power!).

A more complicated, and less familiar, formula arrises in the $n = 3$ case

$$\begin{aligned}\cos 3x + i \sin 3x &= (\cos x + i \sin x)^3 \\ &= \cos^3 x + 3i\cos^2 x \sin x - 3\cos x \sin^2 x - i\sin^3 x \\ &= \left(\cos^3 x - 3\cos x \sin^2 x\right) + i\left(3\cos^2 x \sin x - \sin^3 x\right).\end{aligned}$$

Equating real and imaginary parts via Exercise 3.13

$$\cos 3x = \cos^3 x - 3\cos x \sin^2 x \quad \text{and} \quad \sin 3x = 3\cos^2 x \sin x - \sin^3 x.$$

If we wish to take this an extra step and plug in $\sin^2 x = 1 - \cos^2 x$ we can write $\cos 3x$ as a polynomial in $\cos x$.

$$\cos 3x = 4\cos^3 x - 3\cos x.$$

EXERCISE 3.15. Write $\cos 9x$ as a polynomial in $\cos x$.
Hint! To save time, first write it as a polynomial in $\cos 3x$.
SOLUTION

EXERCISE 3.16. Write $\cos 5x$ as a polynomial in $\cos x$.
Hint! 5 isn't that large of a number, but to save time you could use the fact, derived from the double angle formulas, that

$$\cos 5x = \cos 3x \cos 2x - \sin 3x \sin 2x$$

and previous results, or if you prefer

$$\cos 5x = \cos 4x \cos x - \sin 4x \sin x.$$

SOLUTION

EXERCISE 3.17. Prove that

$$\cos\frac{\pi}{4} = \frac{\sqrt{2}}{2} \qquad \cos\frac{\pi}{8} = \frac{\sqrt{2+\sqrt{2}}}{2} \qquad \cos\frac{\pi}{16} = \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2} \qquad \cos\frac{\pi}{32} = \frac{\sqrt{2+\sqrt{2+\sqrt{2+\sqrt{2}}}}}{2} \ldots$$

SOLUTION

Hint! A simple induction using $\cos 2x = 2\cos^2 x - 1$.

EXERCISE 3.18. Prove that

$$\sin\frac{\pi}{4} = \frac{\sqrt{2}}{2} \qquad \sin\frac{\pi}{8} = \frac{\sqrt{2-\sqrt{2}}}{2} \qquad \sin\frac{\pi}{16} = \frac{\sqrt{2-\sqrt{2+\sqrt{2}}}}{2} \qquad \sin\frac{\pi}{32} = \frac{\sqrt{2-\sqrt{2+\sqrt{2+\sqrt{2}}}}}{2} \ldots$$

SOLUTION

Hint! Another simple induction using $\sin 2x = 2\sin x \cos x$.

3.2.2 Regular Polygons

Complex numbers also give us a handy dandy way of expressing the (a) regular polygon in $\mathbb{R}^2 = \mathbb{C}$.

DEFINITION 3.56. The vertices of the *general regular n–gon*[VIII], is the polygon with vertices

$$\{\{n\}\} \stackrel{\text{def}}{=} \left\{ e^{\frac{2\pi i m}{n}} : m \in \mathbb{Z}, 0 \leq m < n \right\} = \left\{ \cos\frac{2\pi m}{n} + \sin\frac{2\pi m}{n}\, i : m \in \mathbb{Z}, 0 \leq m < n \right\}.$$

(We use the notation $\{\{n\}\}$ to denote either the vertices of a regular n–gon or n–gon itself.)

Since every element of $\{\{n\}\}$ can be expressed as $e^{i\cdot\text{something}}$ we have $\{\{n\}\} \subseteq \mathbb{S}^1$.

EXAMPLE 3.57.

(1) For instance, $\{\{4\}\}$ is a square and it has vertices

$$\{1, i, -1, -i\}.$$

We typically denote this square $B\left(\blacklozenge^3\right)$ instead of $\{\{4\}\}$.

(2) Also note $\{\{3\}\}$ is a (equilateral) triangle and it has vertices

$$\{\{3\}\} = \left\{1, \frac{-1}{2} + \frac{\sqrt{3}}{2}i, -\frac{1}{2} - \frac{\sqrt{3}}{2}i\right\}.$$

We typically denote this triangle $B\left(\blacktriangleright^3\right)$ instead of $\{\{3\}\}$.

(3) The vertices of $\{\{6\}\}$ are

$$\begin{aligned}
e^{\frac{0\pi}{6}i} &= 1\\
e^{\frac{2\pi}{6}i} &= \cos\frac{\pi}{3} + \sin\frac{\pi}{3}\,i = \frac{1}{2} + \frac{\sqrt{3}}{2}i\\
e^{\frac{4\pi}{6}i} &= \cos\frac{2\pi}{3} + \sin\frac{2\pi}{3}\,i = -\frac{1}{2} + \frac{\sqrt{3}}{2}i\\
e^{\frac{6\pi}{6}i} &= \cos\frac{3\pi}{3} + \sin\frac{3\pi}{3}\,i = -1\\
e^{\frac{8\pi}{6}i} &= \cos\frac{4\pi}{3} + \sin\frac{4\pi}{3}\,i = -\frac{1}{2} - \frac{\sqrt{3}}{2}i\\
e^{\frac{10\pi}{6}i} &= \cos\frac{5\pi}{3} + \sin\frac{5\pi}{3}\,i = \frac{1}{2} - \frac{\sqrt{3}}{2}i.
\end{aligned}$$

Since we have that

$$\begin{aligned}
\left(e^{\frac{2\pi im}{n}}\right)^n &= e^{\frac{2\pi imn}{n}}\\
&= e^{2\pi im}\\
&= \cos 2\pi m + \sin 2\pi mi\\
&= 1 + 0\cdot i\\
&= 1
\end{aligned}$$

every element of $\{\{n\}\}$ satisfies the equation $z^n = 1$. On the other hand, since a degree n polynomial contains at most n roots and since $\{\{n\}\}$ contains exactly n elements, the two sets are the same. That is we have proven that

THEOREM 3.58. $\{\{n\}\} = \{z \in \mathbb{C} : z^n = 1\}$. ■

This result is a nice result since it connects geometry (the vertices of a regular n-gon) with algebra (the roots of a polynomial). Based on the last equation, the set $\{\{n\}\}$ is often called the n^{th} *roots of unity* (or one).

DEFINITION 3.59. For any positive integer n we define

$$\chi_n = e^{\frac{2\pi i}{n}}.$$

Notice that $\chi_n \in \{\{n\}\}$. In fact it's the element of the set (besides 1) with the smallest polar angle.

EXAMPLE 3.60. $\chi_4 = i$ and $\chi_6 = \frac{1}{2} + \frac{\sqrt{3}}{2}\,i$.

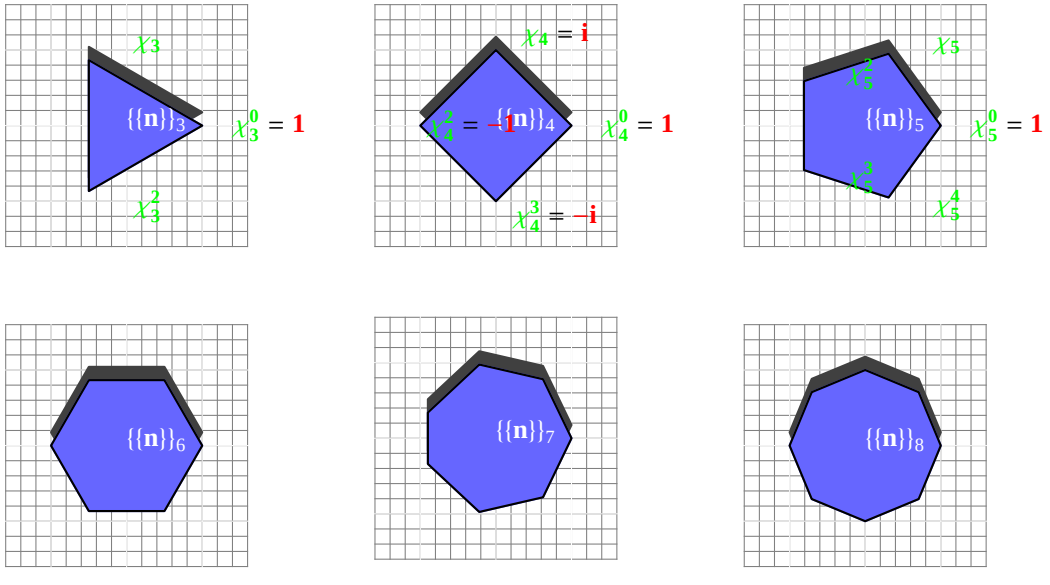


Figure 3.1: Some regular polygons.

EXERCISE 3.19. Suppose that z and w ∈ {{*n*}} (should say they are elements of the vertices of {{*n*}}). Then show that both

(1) $z \cdot w$ **(2)** z^{-1}

are also in {{*n*}}.

SOLUTION

EXERCISE 3.20. Show that every vertex of {{*n*}} is a power between 1 and *n* of χ_n.

SOLUTION

As you are probably familiar with:

$$\cos 0 = 1,\ \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2},\ \cos\frac{\pi}{4} = \frac{\sqrt{2}}{2},\ \cos\frac{\pi}{3} = \frac{1}{2},\ \text{and}\ \cos\frac{\pi}{2} = 0.$$

Let's use these formulas to derive the following useful fact, which compared with the others above is often ignored:

$$\cos\frac{\pi}{5} = \frac{1}{4} + \frac{\sqrt{5}}{4}. \tag{10}$$

The number

$$\frac{1}{2}+\frac{\sqrt{5}}{2}\approx 1.6180333988\ldots$$ [IX]

comes up enough in mathematics to be granted its own name and special symbol.

DEFINITION 3.61. The number

$$\frac{1}{2}+\frac{\sqrt{5}}{2}$$

is called the *golden ratio* and it is denoted by the greek letter ϕ. [X]

So we are trying to show that

$$\cos\frac{\pi}{5}=\frac{\phi}{2}.$$

To this end we expand both sides of DeMoivre's theorem we see

$$-1=e^{\pi i}=\left(e^{\pi i/5}\right)^5=\left(\cos\frac{\pi}{5}+\sin\frac{\pi}{5}\,i\right)^5$$

and compare imaginary parts we see that (need to use Pythogorian theorem,

$$1=e^{ix}\cdot e^{-ix}=(\cos x+\sin x\,i)\cdot(\cos x-\sin x\,i)=\cos^2 x+\sin^2 x$$

as well).

$$0=16\,\sin\left(\frac{\pi}{5}\right)\cos^4\left(\frac{\pi}{5}\right)-12\,\sin\left(\frac{\pi}{5}\right)\cos^2\left(\frac{\pi}{5}\right)+\sin\left(\frac{\pi}{5}\right)$$

and after canceling the non-zero $\sin\frac{\pi}{5}$, and relabeling $x=\cos\frac{\pi}{5}$ we arrive at the equation:

$$0=16x^4-12x^2+1=(4x^2-2x-1)(4x^2+2x-1).$$

The four roots of this equation correspond to the four values

$$\cos\frac{\pi}{5},\ \cos\frac{2\pi}{5},\ \cos\frac{3\pi}{5},\ \cos\frac{4\pi}{5}$$

(since the same calculation applies). Now using the quadratic formula on both factors shows that the four roots are

$$\pm\frac{2}{8}\pm\frac{\sqrt{4+4\cdot 4\cdot 1}}{8}=\pm\frac{1}{4}\pm\frac{\sqrt{5}}{4}.$$

Since cosine is a decreasing function from 0 to π we know that $\cos\frac{\pi}{5}$ is the largest of these four roots. That is

$$\cos\frac{\pi}{5}=\frac{1}{4}+\frac{\sqrt{5}}{4}=\frac{\phi}{2}$$

[IX] The only thing we need to know about the numerical value of the golden ratio is that it's greater than 1. Even that we only use once for one of the exercises!

[X] We make a distinction between the Greek letter ϕ and its "cursive" version φ. The former will always denote the golden ratio while the latter will never mean the golden ratio.

as promised.

Let's talk a bit more about the golden ration, ϕ. For many purposes, it's important to note that the golden ratio, ϕ, which we defined to be

$$2\cos\frac{\pi}{5} = \frac{1}{2} + \frac{\sqrt{5}}{2},$$

is a root of the polynomial

$$x^2 - x - 1. \tag{11}$$

And, as such, it satisfies

$$\phi^2 = \phi + 1. \tag{12}$$

(In words this says that squaring ϕ is the same as adding 1 to it, a remarkable relation indeed!) This is obtained by plugging $\frac{\phi}{2} = x$ into the equation $4x^2 = 2x + 1$, or, more directly, from the definition of ϕ as $\phi = \frac{1}{2} + \frac{\sqrt{5}}{2}$.[XI]

DEFINITION 3.62. The *Fibonacci sequence* is defined recursively as follows:

$$F_1 = 1,\ F_2 = 1 \text{ and for } n > 2,\ F_n \stackrel{\text{def}}{=} F_{n-1} + F_{n-2}.$$

So the first few terms of the sequence are 1, 1, 2, 3, 5, 8, 13, 21, 34, 55. See Appendix K for more terms in this sequence.

EXERCISE 3.21. Prove that the golden ration, ϕ, satisfies

$$\phi = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}}} \tag{13}$$

SOLUTION

EXERCISE 3.22. Suppose that Φ is something (could be a number, or even a matrix) which satisfies $\Phi^2 = \Phi + 1$ [XII] Prove that

$$\Phi^n = F_n\Phi + F_{n-1}.$$

SOLUTION

Hint! Simply use induction on n.

[XI] Since this means that

$$2\phi - 1 = \sqrt{5} \Rightarrow (2\phi - 1)^2 = 5 \Rightarrow 4\phi^2 - 4\phi + 1 = 5 \Rightarrow \phi^2 = \phi + 1.$$

Let $\widetilde{\phi}$ as $1-\phi$. I claim that this the other root (besides the golden ratio) of the polynomial $x^2 = x+1$. To see we do the following straight forward calculation:

$$\begin{aligned}\widetilde{\phi}^2 - \widetilde{\phi} - 1 &= (1-\phi)^2 - (1-\phi) - 1 \\ &= (1 - 2\phi + \phi^2) - (1-\phi) - 1 \\ &= (1 - 2\phi + [\phi + 1]) - (1-\phi) - 1 \quad \text{(Using (12))} \\ &= (2-\phi) - 1 + \phi - 1 \\ &= 0.\end{aligned}$$

EXERCISE 3.23.

SOLUTION

(1) Prove for all positive integers that

$$F_n = \frac{\phi^n - \widetilde{\phi}^n}{\phi - \widetilde{\phi}}. \tag{14}$$

where F_n is the n^{th} term in he Fibonacci sequence (defined in the previous exercise) and ϕ is the golden ratio and its conjugate, $\widetilde{\phi}$.
Hint! There is no need for induction in this proof.

(2) Use `MAPLE` or another computer program and the above formula, (14), to compute F_{101} and F_{304}.

Hint! F_{304} has 64 digits!(only 64 digits ...)

Note that since

$$\widetilde{\phi} \cdot \phi = (1-\phi)\cdot\phi = \phi - \phi^2 = \phi - 1 - \phi = -1^{\text{XIII}}$$

we have that

$$\phi^{-1} = -\widetilde{\phi} = \phi - 1.$$

EXERCISE 3.24. Prove that

SOLUTION

$$\widetilde{\phi}^n = (F_{n+1} - \phi \cdot F_n)$$

and

$$\phi^{-n} = (-1)^n \cdot (F_{n+1} - \phi \cdot F_n).$$

[XII] For example by (12) the golden ratio defined above satisfies this. More exotically, the "golden matrix"

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

does as well.

[XIII] Alternatively, we knew this already since ϕ and $\widetilde{\phi}$ are the two roots $x^2 - x - 1$ and the product of the two roots of any quadradic equation always equals the last coefficient, in this case -1.

Exercise 3.25. Prove that

$$\phi^n + \phi^{-n} = \begin{cases} F_{n+1} + F_{n-1} & \text{if } n \text{ is even} \\ \sqrt{5} \cdot F_n & \text{if } n \text{ is odd.} \end{cases}$$

SOLUTION

Exercise 3.26. Prove that the ratio of consecutive terms in Fibonacci sequence approaches ϕ. That is prove that

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \phi.$$

SOLUTION

Hint! Since $\left|\widetilde{\phi}\right| < 1$ we know that the limit $\widetilde{\phi}^n$ as $n \to \infty$ is 0. Use this and Exercise **??** or Exercise **??**.

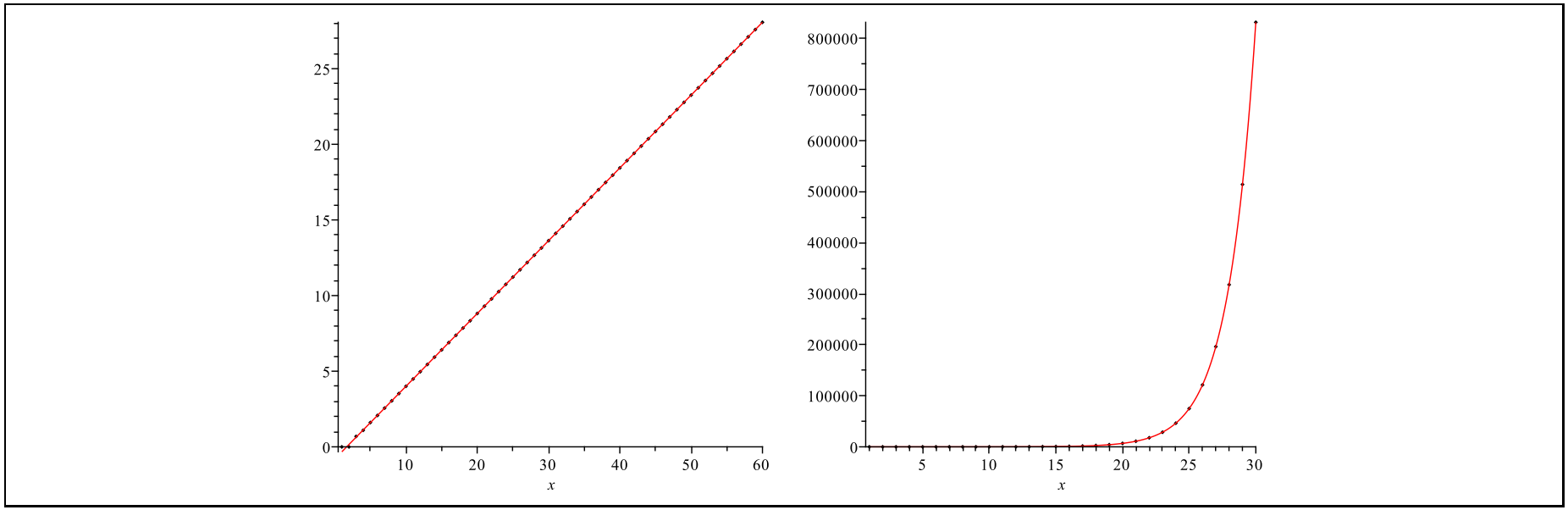


Figure 3.2: Here, in the left picture, the dotted line is $\ln F_n$ (shown as dots) plotted against $\ln(\phi) \cdot x - \ln\sqrt{5} \approx 0.4812x - 0.804$ (shown by the red line). The second picture shows the plots of F_n (as dots) and $\phi^n/\sqrt{5}$ for $n \leq 30$ (as the red line). (Stuff gets too big too fast after this point so the log plots are better for larger values.) These graphs show how close the approximation $F_n \approx \frac{\phi^n}{\sqrt{5}}$ is.

3.3 QUATERNIONS

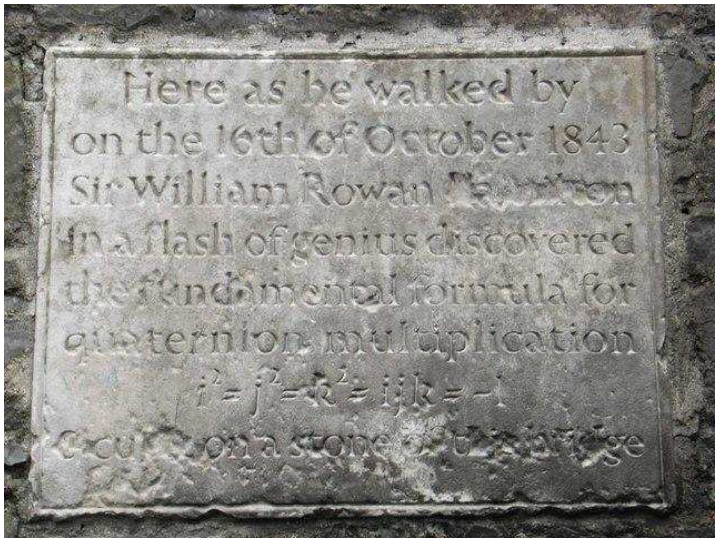


Figure 3.3: The equations read $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$.

If you are not up on your cross and dot products you should check out Appendix C which starts on page 741.

3.3.1 BASIC DEFINITIONS AND THEOREMS

Just like we can think of $\mathbb{C}$ as $\mathbb{R}^2$ with a certain addition and multiplication defined on it, we can think of the quaternions $\mathbb{H}$ as $\mathbb{R}^4$, $\mathbb{C}^2$, or even $\mathbb{R} \times \mathbb{R}^3$ with fancy multiplications defined on them.

{need to put some blah blah blah about quaternions here}

With the exception of proving associativity, in my opinion, the properties of quaternions are most readily seen by using the $\mathbb{R} \times \mathbb{R}^3$ approach. This approach uses the most vector calculus. Typically students prefer the $\mathbb{R}^4$ approach since it's easier to do things out the long way. Unfortunately this approach makes it very tedious to derive the basic properties of the quaternions. The $\mathbb{C} \times \mathbb{C}$ approach is the easiest way of proving that the multiplication of quaternions is associative. Also the multiplication looks nearly the same on $\mathbb{C} \times \mathbb{C}$ as it did back in Section 3.2 to define the multiplication on $\mathbb{C}$ when viewing this as $\mathbb{R} \times \mathbb{R}$.

DEFINITION 3.63. Define $\mathbb{H} \stackrel{\text{def}}{=} \mathbb{R} \times \mathbb{R}^3$. Elements of $\mathbb{H}$ are known as *quaternions* and $\mathbb{H}$ is called *the quaternions*.

That is a quaternion is a number followed by a vector with three coordinates:

$$(a, \mathbf{v}) \text{ where } a \in \mathbb{R} \text{ and } \mathbf{v} \in \mathbb{R}^3 .$$

EXAMPLE 3.64. So quaternions are things that look like $(5, \mathbf{i} + 3\mathbf{j} + \mathbf{k})$, $(6, 2\mathbf{j} + \mathbf{k})$, etc.

In practice, breaking things up this way (a single followed by a triple) is only a matter of taste:

Convention 3.65. We will not rigorously stick to thinking of a quaternion as an element of $\mathbb{R} \times \mathbb{R}^3$, i.e. a number following by a triple of numbers. Often we will think of it as an element of $\mathbb{R}^4$, i.e. just four numbers given at once. Occasionally, we will also think of a quaternion as two pairs (actually complex numbers) of numbers.

This is related to what we blabbed about in Remark 2.10 back on page 9.

We also think of $\mathbb{R}$ and $\mathbb{R}^3$ as literally being subsets of $\mathbb{H}$ in the obvious ways:

Convention 3.66.

(1) We identify a quaternion of the form $(c, \mathbf{0})$ with the real number c and the real numbers, $\mathbb{R}$, with the set of quaternions of this form.

(2) We identify a quaternion of the form $(0, \mathbf{v})$ with the 3-vector $\mathbf{v}$ and $\mathbb{R}^3$ with the set of quaternions of this form.

$c \leftrightsquigarrow (c, \mathbf{0})$ AND $\mathbf{v} \leftrightsquigarrow (0, \mathbf{v})$.

This should seem reasonable: this is what we typically do anyway, we think of the real numbers as being subset of the plane, typically as the set of points in the plane on the x-axis. Note that with this convention we have that $\mathbb{R}$ and $\mathbb{R}^3$ meet only at $(0, \mathbf{0})$ the origin.

Definition 3.67. For a quaternion $\mathsf{q} = (a, \mathbf{v})$

(1) The number a is the *real part* of q. We denote this as $\mathbf{Re}\,\mathsf{q} = a$.

(2) The vector $\mathbf{v}$ is the *pure part* of q. We denote this as $\mathbf{Pu}\,\mathsf{q} = \mathbf{v}$.

In several other texts what we call the pure part of a quaternion is called the *vector part*.

Definition 3.68. A quaternion q is

(1) *real* if it is equal to its real part, i.e. $\mathbf{Re}\,\mathsf{q} = \mathsf{q}$.

(2) *pure* if it is equal to its pure part, i.e. $\mathbf{Pu}\,\mathsf{q} = \mathsf{q}$

As such a quaternion is pure iff it's real part is zero and real if its pure part is zero. Of course the only quaternion which is both pure and real is the zero quaternion: $(0, \mathbf{0})$. Moreover note that the pure part of any quaternion is a pure quaternion and the real part of any quaternion is a real quaternion. Of course most quaternions are neither real nor pure as the next example illustrates.

Example 3.69.

(1) Let

$$\mathsf{q} = (5, \mathbf{i} + 3\mathbf{j} + \mathbf{k}).$$

Then the real part of q, $\mathbf{Re}\,\mathsf{q}$, is the real number 5, and the pure part of q, $\mathbf{Pu}\,\mathsf{q}$, is the vector $\mathbf{i} + 3\mathbf{j} + \mathbf{k}$. Note that q is neither real nor pure.

(2) The quaternion $\mathsf{q} := (0, 2\mathbf{i} + \mathbf{k})$ is pure. It's real part, $\mathbf{Re}\,\mathsf{q}$, is 0 and its pure part, $\mathbf{Pu}\,\mathsf{q}$, is $2\mathbf{i} + \mathbf{k}$. Under the convention, Convention 3.66, we could write $\mathsf{q} = 2\mathbf{i} + \mathbf{k}$.

(3) The quaternion $\mathsf{q} = (5, \mathbf{0})$ is real. It's real part, $\mathbf{Re}\,\mathsf{q}$, is the number 5, its pure part, $\mathbf{Pu}\,\mathsf{q}$, is the zero vector, $\mathbf{0}$. Under the convention, Convention 3.66, we could write $\mathsf{q} = 5$.

WARNING! 3.70. Do not attempt to visualize a quaternion at this point! These dudes are 4-dimensional. We will later see that quaternions are the best way to represent a rotation in 3-dimensions. If you are desperate to at least see something, instead of thinking of quaternions as a dude in 4 dimensions, you can think of them as a dude in 3 dimensions (his position w.r.t a reasonable coordinate system set up at say, his house, corresponding to the pure part of a quaternions) with a device to measure (say) the amount of time (in say seconds) since he started his watch (which corresponds to the real part of a quaternion). In this way a pure quaternion corresponds to our dude have not having started his watch and a real quaternion corresponds to our dude being lazy and not leaving home.

Addition is defined for quaternions in the obvious way –by adding coordinates. That is for quaternions $(a, \mathbf{v})$ and $(b, \mathbf{w}) \in \mathbb{H}$.

$$(a, \mathbf{v}) + (b, \mathbf{w}) = (a + b, \mathbf{v} + \mathbf{w})$$

Also scalar multiplication of a quaternion $(a, \mathbf{v})$ by a real number r is also defined in the obvious way:

$$r \cdot (a, \mathbf{v}) \stackrel{\text{def}}{=} (r \cdot a, r \cdot \mathbf{v}).$$

In even more detail these operations are

$$(a, (v_1, v_2, v_3)) + (b, (w_1, w_2, w_3)) = (a + b, (v_1 + w_1, v_2 + w_2 + v_3 + w_3))$$

and

$$r \cdot (a, (v_1, v_2, v_3)) = (r \cdot a, (r \cdot v_1, r \cdot v_2, r \cdot v_3)).$$

Note that both of these operations are exactly the same as the usual, everyday operations when we think of a quaternion as an element of $\mathbb{R}^4$ as in Convention 3.65. This also means that there is no need to prove any of the basic properties Proposition **??** on page **??** again in our new situation. However, of course, if we were to prove these properties again it would be just as trivial as it was originally!

Note with these definitions and our conventions, we have for any quaternion $\mathsf{q} = (a, \mathbf{v})$, we have

$$\mathsf{q} = a + \mathbf{v}.$$

You should try to take note by the bolding which is a number and which is a vector since it is also true that

$$\mathsf{q} = \mathbf{v} + a$$

as well. This also means that q is the sum of its real and pure parts:

$$\mathsf{q} = \mathbf{Re}(\mathsf{q}) + \mathbf{Pu}(\mathsf{q}).$$

Remark 3.71. Another advantage of thinking of a quaternion as a four tuple is that it is now obvious how to define the length of a quaternion q, q. We will do this more in content in Definition 3.84 in a bit. as well as take their lengths, see Definition 3.84 below.

Despite the advantages of thinking of $\mathbb{H}$ as four numbers given at once, the biggest advantage of breaking a quaternion down into a number and 3-vector is that it makes the multiplication, which we have yet to define, of two quaternions have a more intuitive feeling. (However you may not feel that way initially, but I hope you will in a page or two.)

Definition 3.72. ($\mathbb{R} \times \mathbb{R}^3$ Version) Given quaternions $\mathsf{q} = a + \mathbf{v}$ and $\mathsf{p} = b + \mathbf{w}$ we define *the* (quaternion) multiplication $*$ by the formula

$$\mathsf{q} * \mathsf{p} \stackrel{\text{def}}{=} (ab - \mathbf{v} \bullet \mathbf{w}, a\mathbf{w} + b\mathbf{v} + \mathbf{v} \times \mathbf{w})$$

where $\bullet$ is the dot product and $\times$ is the cross product of the vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$. [1]

Written as a sum, the product in the definition is therefore

$$ab - \mathbf{v} \bullet \mathbf{w} + a\mathbf{w} + b\mathbf{v} + \mathbf{v} \times \mathbf{w}.$$

The (slight) disadvantage of doing this is that in order to keep things straight you cannot forget that the dot product of two vectors is a number and the cross product of two 3-vectors is another 3-vector, so in this case parenthesis are especially useful:

$$(ab - \mathbf{v} \bullet \mathbf{w}) + (a\mathbf{w} + b\mathbf{v} + \mathbf{v} \times \mathbf{w}).$$

Note there seems to be some ambiguity for real numbers when talking about the product: do we mean scalar multiplication or quaternion multiplication? But luckily there's no distinction since

$$\begin{aligned} r * (a, \mathbf{v}) &= (r, \mathbf{0}) * (a, \mathbf{v}) \\ &= (ra - \mathbf{0} \bullet \mathbf{v}, r\mathbf{v} + a\mathbf{0} + \mathbf{0} \times \mathbf{v}) \\ &= (ra, r\mathbf{v}) \\ &= r \cdot (a, \mathbf{v}). \end{aligned}$$

This definition (at least initially) seems rather complicated. The reason for the complicated definition is that it turns out to be (essentially) the only definition one can define on $\mathbb{H}$ for $*$ so that we have some nice properties. For instance, see Exercise 4.16 coming up on page 229 for more details on this. An example of an awesome property that will hold will be seen in Theorem 3.98: namely that $*$ is associative. This is despite the fact that neither cross product nor dot product is associative.

Example 3.73.

$$\begin{aligned} (1 + \mathbf{i} + \mathbf{j}) * (2 + \mathbf{i} + \mathbf{k}) &= \Big(2 - (\mathbf{i} + \mathbf{j}) \bullet (\mathbf{i} + \mathbf{k})\Big) + \Big((\mathbf{i} + \mathbf{k}) + 2 \cdot (\mathbf{i} + \mathbf{j}) + (\mathbf{i} + \mathbf{j}) \times (\mathbf{i} + \mathbf{k})\Big) \\ &= (2 - 1) + (3\mathbf{i} + 2\mathbf{j} + \mathbf{k} + (-\mathbf{j} - \mathbf{k} + \mathbf{i})) \\ &= 1 + 4\mathbf{i} + \mathbf{j}. \end{aligned}$$

[1] We will eventually drop the $*$ and simply write multiplication simply as concatenation or as $\cdot$. It is for this reason we exaggerate the dot in the dot product by using a HUGE dot!!

Example 3.74. Let $q_1 = 1 + \mathbf{i}$ and $q_2 = 2 + \mathbf{i} + \mathbf{j}$. Then we have that

$$q_1 + q_2 = 3 + 2\mathbf{i} + \mathbf{j},$$

$$\begin{aligned} q_1 * q_2 &= (1 + \mathbf{i}) * (2 + \mathbf{i} + \mathbf{j}) \\ &= (1 \cdot 2 - \mathbf{i} \bullet (\mathbf{i} + \mathbf{j})) + (1 \cdot (\mathbf{i} + \mathbf{j}) + 2 \cdot \mathbf{i} + \mathbf{i} \times (\mathbf{i} + \mathbf{j})) \\ &= (2 - 1) + (\mathbf{i} + \mathbf{j} + 2\mathbf{i}) + (\mathbf{k}) \\ &= 1 + 3\mathbf{i} + \mathbf{j} + \mathbf{k} \end{aligned}$$

and

$$\begin{aligned} q_2 * q_1 &= (2 + \mathbf{i} + \mathbf{j}) * (1, \mathbf{i}) \\ &= (2 \cdot 1 - (\mathbf{i} + \mathbf{j}) \bullet \mathbf{i}) + (2 \cdot \mathbf{i} + 1 \cdot (\mathbf{i} + \mathbf{j}) + (\mathbf{i} + \mathbf{j}) \times \mathbf{i})) \\ &= (2 - 1) + (\mathbf{i} + \mathbf{j} + 2\mathbf{i}) + (-\mathbf{k}) \\ &= 1 + 3\mathbf{i} + \mathbf{j} - \mathbf{k}. \end{aligned}$$

Notice that in particular that $q_1 * q_2 \neq q_2 * q_1$. In other words q_1 and q_2 do not commute, and the operation, $*$, is not commutative.

Exercise 3.27. Calculate $(2 + \mathbf{i} + \mathbf{j}) * (1 + \mathbf{i} + \mathbf{j} + \mathbf{k})$ using only the present definitions.

Solution

Exercise 3.28. Prove that for any two quaternions q and p we have

$$q * p - p * q = 2 \cdot (\mathbf{Pu}\, q \times \mathbf{Pu}\, p)\,.$$

Solution

Using the fact that two vectors are parallel, i.e. iff their cross product is $\mathbf{0}$, the previous exercise proves the following theorem:

Theorem 3.75. *Two quaternions* q *and* p *commute* (*meaning* $p * q = q * p$) *iff their pure parts are parallel:*

$$\mathbf{Pu}\, q \asymp \mathbf{Pu}\, p.$$

Proof. Two quaternions q and p commute iff

$$q * p - p * q = \mathbf{0}.$$

But the exercise, Exercise 3.28, says that this is equal to twice the cross product of the two pure parts, so

$$\mathbf{0} = 2\,(\mathbf{Pu}\, q \times \mathbf{Pu}\, p) \Rightarrow \mathbf{0} = \mathbf{Pu}\, q \times \mathbf{Pu}\, p.$$

Since the cross product of two vectors is $\mathbf{0}$ iff the two vectors we start with are parallel[II], this means that two quaternions ■

For example, we can instantly see that of the three quaternions:

$$3 + 2\mathbf{i} + 5\mathbf{j} + \mathbf{k} \quad \text{and} \quad 1 + 2\mathbf{i} + 6\mathbf{k} \quad \text{and} \quad 1 - 4\mathbf{i} - 10\mathbf{j} - 2\mathbf{k}$$

only the first and third commute. Moreover, since the zero vector is parallel to every vector this theorem also says that every real quaternion commutes with every quaternion. (Of course we already knew that since we've already shown that quaternion multiplication by a number is the same as scalar multiplication by that same number.)

Remark 3.76. Let's point out some quick facts about the multiplication. For starters, we have that for $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$ and for c and $c' \in \mathbb{R}$, that

$$\mathbf{v} * \mathbf{w} = (0 \cdot 0 - \mathbf{v} \bullet \mathbf{w}) + (0 \cdot \mathbf{w} + 0 \cdot \mathbf{v} + \mathbf{v} \times \mathbf{w}) = -\mathbf{v} \bullet \mathbf{w} + \mathbf{v} \times \mathbf{w} \tag{15}$$

and

$$c * (a, \mathbf{v}) = c \cdot (a, \mathbf{v}) = (ca, c\mathbf{v}) = (a, \mathbf{v}) * c \tag{16}$$

In particular, for any quaternion q we have that $(1, \mathbf{0}) * \mathsf{q} = \mathsf{q} = \mathsf{q} * (1, \mathbf{0})$. In this way $(1, \mathbf{0})$ is a multiplicative identity. Furthermore, we have that multiplication of real quaternions is the same as the corresponding multiplication of real numbers.

Example 3.77. For the standard unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ we have that each of their products is equal to -1 (... ok, $(-1, \mathbf{0})$) whenever they are the same and equal to their cross product whenever they are different. For instance:

$$\mathbf{i} * \mathbf{i} = -\mathbf{i} \bullet \mathbf{i} = -1 \qquad \mathbf{i} * \mathbf{j} = \mathbf{i} \times \mathbf{j} = \mathbf{k}$$
$$\mathbf{j} * \mathbf{j} = -\mathbf{j} \bullet \mathbf{j} + \mathbf{j} \times \mathbf{j} = -1 \qquad \mathbf{i} * \mathbf{k} = \mathbf{i} \times \mathbf{k} = -\mathbf{j}.$$

When we wish to be perfectly clear whether or not we wish to discuss a vector such as $\mathbf{i}$ as a quaternion we unbold it. More precisely, we are defining the three special quaternions:

$$i = (0, \mathbf{i}) = \left(0, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right), \qquad j = (0, \mathbf{j}) = \left(0, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right), \quad \text{and} \quad k = 0, \mathbf{k}) = \left(0, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right).$$

The reason for doing this is because not only does $\mathbf{i}$ behave like the vector it denotes, but since as we just pointed out $\mathbf{i} * \mathbf{i} = \mathbf{i}^2 = -1$, it also behaves like our favorite complex number, i. The distinction between using $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ or i, j, and k is made only to highlight either the geometric (i.e. when talking about cross products, lengths, dot products etc.) or algebraic nature (i.e. when talking about powers, inverses, and orders) respectively. However, quite often, we use the bolded versions for no deep reason other than the fact that bold letters pop.

Theorem 3.78. *Quaternion multiplication is bilinear. That is given any three quaternions* q, q', *and* q'' *and any real number* c *we have*

(Left Distributive)	$\mathsf{q} * (\mathsf{q}' + \mathsf{q}'') = \mathsf{q} * \mathsf{q}' + \mathsf{q} * \mathsf{q}''$
(Right Distributive)	$(\mathsf{q} + \mathsf{q}') * \mathsf{q}'' = \mathsf{q} * \mathsf{q}'' + \mathsf{q}' * \mathsf{q}''$
(Scalar Pulling)	$c \cdot (\mathsf{q} * \mathsf{q}') = (c \cdot \mathsf{q}) * \mathsf{q}' = \mathsf{q} * (c \cdot \mathsf{q}')$.

[II] This is a well known result from vector calculus and it makes perfectly good geometric sense and a formal proof based on what we already know can be given based on properties of the determinant. We give a different proof of this fact later in Corollary 3.102**(1)**.

Proof. We prove only right distributivity (leaving left distributivity and scalar pulling to you). Let $\mathfrak{q} = q_R + \mathbf{q}_P$, $\mathfrak{q}' = q'_R + \mathbf{q}'_P$ and $\mathfrak{q}'' = q''_R + \mathbf{q}''_P$. Then

$$\begin{aligned}
(\mathfrak{q} + \mathfrak{q}') * \mathfrak{q}'' &= \left(\overset{\mathfrak{q}}{\overset{\shortparallel}{q_R + \mathbf{q}_P}} + \overset{\mathfrak{q}'}{\overset{\shortparallel}{q'_R + \mathbf{q}'_P}} \right) * \overset{\mathfrak{q}''}{\overset{\shortparallel}{q''_R + \mathbf{q}''_P}} \\
&= \Big((q_R + q'_R) + (\mathbf{q}_P + \mathbf{q}'_P) \Big) * (q''_R + \mathbf{q}''_P) \\
&= \Big((q_R + q'_R) \cdot q''_R - (\mathbf{q}_P + \mathbf{q}'_P) \bullet \mathbf{q}''_P \Big) + \Big((q_R + q'_R) \cdot \mathbf{q}''_P + q''_R \cdot (\mathbf{q}_P + \mathbf{q}'_P) + (\mathbf{q}_P + \mathbf{q}'_P) \times \mathbf{q}''_P \Big) \\
&= \Big((q_R \cdot q''_R + q'_R \cdot q''_R) - (\mathbf{q}_P \bullet \mathbf{q}''_P + \mathbf{q}_P \bullet \mathbf{q}''_P) \Big) + \Big((q_R \cdot \mathbf{q}''_P + q'_R \cdot \mathbf{q}''_P + q''_R \cdot \mathbf{q}_P + q''_R \cdot \mathbf{q}'_P) + (\mathbf{q}_P \times \mathbf{q}''_P + \mathbf{q}'_P \times \mathbf{q}''_P) \Big)
\end{aligned} \tag{17}$$

and meanwhile

$$\mathfrak{q} * \mathfrak{q}' + \mathfrak{q} * \mathfrak{q}'' = \Big[(q_R \cdot q''_R - \mathbf{q}_P \bullet \mathbf{q}''_P) + (q_R \cdot \mathbf{q}''_P + q''_R \cdot \mathbf{q}_P + \mathbf{q}_P \times \mathbf{q}''_P) \Big] + \Big[(q'_R \cdot q''_R - \mathbf{q}'_P \bullet \mathbf{q}''_P) + (q'_R \cdot \mathbf{q}''_P + q''_R \cdot \mathbf{q}'_P + \mathbf{q}'_P \times \mathbf{q}''_P) \Big]. \tag{18}$$

To get from the second to last to the last lines of (17) we have used the facts that each of the operations

- Regular multiplication of numbers
- Scalar multiplication of a number and a vector
- The dot product of two vectors
- The cross product of two vectors

is both left and right distributive. Now the quantities from Equations (17) and (18) are equal: simply rearrange the terms of the addition from one of the equations to get the other. Therefore the right distributive property holds. [III] ∎

Remark 3.79. A plus of the bilinear property is that we can think of multiplication of quaternions by breaking it down into three possibilities

(1) For two real numbers r_1 and r_2, their quaternion product is just the boring, everyday multiplication of real numbers:

$$r_1 * r_2 = r_1 \cdot r_2 = r_2 \cdot r_1 = r_2 * r_1.$$

[III] I hope you notice that in this proof how we only used the facts that cross, dot, and the usual products of real numbers are bilinear. Perhaps from this, you could conclude that something defined along the lines of something that I completely make up like

$$(q_R + \mathbf{q}_P) \star (q'_R + \mathbf{q}'_P) \overset{\text{def}}{=} (q_R \cdot q'_R + \mathbf{q}_P \bullet \mathbf{q}'_P) + (q_R \cdot \mathbf{q}'_P + q'_R \cdot \mathbf{q}_P + \mathbf{q}_P \times \mathbf{q}'_P)$$

is also bilinear. However, with another of these randomly defined products, we will more than likely lose our precious associative property, and indeed we do. We study this more exotic multiplication in Exercise 4.16 coming up on page 229 in the next chapter.

(2) For a real number r and a pure quaternion $\mathbf{v} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$, their product is the same as scalar multiplication in $\mathbb{R}^3$. That is

$$\begin{aligned} r * \mathbf{v} = r * (a\mathbf{i} + b\mathbf{j} + c\mathbf{k}) &= r(a\mathbf{i} + b\mathbf{j} + c\mathbf{k}) \\ &= ra\mathbf{i} + rb\mathbf{j} + rc\mathbf{k} = (a\mathbf{i} + b\mathbf{j} + c\mathbf{k}) * r = \mathbf{v} * r. \end{aligned}$$

(3) For two pure quaternions, $\mathbf{v}$ and $\mathbf{w}$, their multiplication is the cross product minus the dot product. That is

$$\mathbf{v} * \mathbf{w} = -\mathbf{v} \bullet \mathbf{w} + \mathbf{v} \times \mathbf{w}.$$

In fact, these three properties determine $*$ completely (and make the definition of $*$ seem more digestible). Since using only the three properties we see that

$$\begin{aligned} (a + \mathbf{v}) * (b + \mathbf{w}) &= a * (b + \mathbf{w}) + \mathbf{v} * (b + \mathbf{w}) && \text{(Distrib)} \\ &= a * b + a * \mathbf{w} + \mathbf{v} * b + \mathbf{v} * \mathbf{w} && \text{(Distrib)} \\ &= a \cdot b + a * \mathbf{w} + \mathbf{v} * b + \mathbf{v} * \mathbf{w} && \text{((1) above)} \\ &= a \cdot b + a\mathbf{w} + b\mathbf{v} + \mathbf{v} * \mathbf{w} && \text{((2) above)} \\ &= a \cdot b + a\mathbf{w} + b\mathbf{v} - (\mathbf{v} \bullet \mathbf{w}) + \mathbf{v} \times \mathbf{w} && \text{((3) above)} \end{aligned}$$

which of course was the original definition of the multiplication of quaternions. Thus, we can think of the quaternion multiplication as, sort of, the cheapest way of turning the cross product into an associative operation.

Quaternion multiplication is cross product minus dot product for vectors and scalar multiplication for numbers.

We can further break this down since we know that the biggest constituents of the quaternion product, the dot and the cross product, are also bilinear. So we can multiply out any two quaternions just by knowing how to multiply 1, **i**, **j**, and **k**. Actually, since $1 * \mathsf{q} = \mathsf{q} = \mathsf{q} * 1$ for any quaternion, we really only need to know how to multiply the pure quaternions **i**, **j**, and **k** by each other. Since for pure quaternions, we have that

$$\mathbf{v} * \mathbf{w} = \mathbf{v} \times \mathbf{w} - \mathbf{v} \bullet \mathbf{w}$$

we can simply subtract the table for the dot product, Table A.1 on page 748, from the table for the cross product, Table A.2 on page 751, to get Table 3.1.

Exercise 3.29. Show that the equations

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$$

from Figure 3.3 determine all the information from Table 3.1 and vice versa.

$*$	$\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$
$\mathbf{i}$	-1	$\mathbf{k}$	$-\mathbf{j}$
$\mathbf{j}$	$-\mathbf{k}$	-1	$\mathbf{i}$
$\mathbf{k}$	$\mathbf{j}$	$-\mathbf{i}$	-1

Table 3.1: The quaternion product of **i**, **j**, and **k**.

EXAMPLE 3.80. Using this and distributing out it's possible to evaluate every product (however this method often takes a bit longer). For instance, we can multiply the quaternions

$$\mathsf{q} = 1 + \mathbf{i} + \mathbf{j} \quad \text{and} \quad \mathsf{p} = 1 + 2\mathbf{i} + \mathbf{k}$$

via

$$\begin{aligned}(1+\mathbf{i}+\mathbf{j}) * (1+2\mathbf{i}+\mathbf{k}) &= (1+2\mathbf{i}+\mathbf{k}) + \mathbf{i} * (1+2\mathbf{i}+\mathbf{k}) + \mathbf{j} * (1+2\mathbf{i}+\mathbf{k}) \\ &= (1+2\mathbf{i}+\mathbf{k}) + (\mathbf{i} - 2 - \mathbf{j}) + (\mathbf{j} - 2\mathbf{k} + \mathbf{i}) \\ &= (1-2) + (2+1+1)\mathbf{i} + \cancel{(-1+1)\mathbf{j}} + (1-2)\mathbf{k} \\ &= -1 + 4\mathbf{i} - \mathbf{k}.\end{aligned}$$

In fact using this, or plugging in the formulas for the dot and cross products directly, we can come up with a general formula which, given its seeming lack of pattern, hopefully makes you appreciate the "cross product minus dot product" approach taken so far. Nevertheless, students often feel more comfortable knowing that there is a concrete formula out there to use

THEOREM 3.81. ($\mathbb{R}^4$ Version)

$$\begin{aligned}(a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k}) * (x + y\,\mathbf{i} + z\,\mathbf{j} + w\,\mathbf{k}) \\ = (ax - by - cz - dw) + (bx + ay + dz - cw)\,\mathbf{i} \\ + (cx + az + bw - dy)\,\mathbf{j} + (aw + dx + cy - bz)\,\mathbf{k}.\end{aligned} \tag{19}$$

Students typically like to start out using this method to calculate things, but they quickly learn that proving general formulas is an incredibly boring and is hardly an insightful experience. I call this the $\mathbb{R}^4$ version of quaternion multiplication as opposed to the $\mathbb{R} \times \mathbb{R}^3$ version from Definition 3.72 used originally. Later, in 3.95 we will see a third and final version of quaternion multiplication.

DEFINITION 3.82. The conjugate of a quaternion $\mathsf{q} = (q_R, \mathbf{q}_P)$ is the quaternion $\overline{\mathsf{q}} = (q_R, -\mathbf{q}_P)$.

So the conjugate of a quaternion keeps the real part the same, but it flips the sign of the pure part. In this way it is analogous to the complex conjugate of a complex number.

In more detail

$$\overline{a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}.$$

Note we have

$$\overline{\overline{\mathsf{q}}} = \overline{\overline{(q_R, \mathbf{q}_P)}} = \overline{(q_R, -\mathbf{q}_P)} = (q_R, \mathbf{q}_P) = \mathsf{q}.$$

That is conjugation is an involution (something that you do twice and end up doing nothing).[IV]

[IV] This is even more obvious in words since if you flip the sign of the pure part twice you'll end up doing nothing.

EXAMPLE 3.83. $\overline{5 + 3\mathbf{i} + 5\mathbf{j}} = 5 - 3\mathbf{i} - 5\mathbf{j}$.

EXERCISE 3.30. Prove that for any $\mathsf{q} \in \mathbb{H}$

$$\mathbf{Re}\,\mathsf{q} = \frac{1}{2}(\mathsf{q} + \overline{\mathsf{q}})$$

(the real part of q) and

$$\mathbf{Pu}\,\mathsf{q} = \frac{1}{2}(\mathsf{q} - \overline{\mathsf{q}})$$

(the pure part of q).

SOLUTION

EXERCISE 3.31. Prove that conjugation is linear, i.e. show that for any two quaternions q_1 and q_2 and any real number c that

$$\overline{\mathsf{q} + \mathsf{p}} = \overline{\mathsf{q}} + \overline{\mathsf{p}} \quad \text{and} \quad \overline{c \cdot \mathsf{q}} = c\overline{\mathsf{q}}.$$

SOLUTION

EXERCISE 3.32. Show for any quaternion q that

$$\overline{\mathsf{q}} = -\frac{1}{2}(\mathsf{q} + \mathbf{i}\mathsf{q}\mathbf{i} + \mathbf{j}\mathsf{q}\mathbf{j} + \mathbf{k}\mathsf{q}\mathbf{k}).$$

SOLUTION

As already pointed out, it's not a coincidence that we decided to denote the quaternion $(0, \mathbf{i})$ in the same way that we denoted everybody's favorite imaginary friend, i. We can exaggerate this coincidence even further by noting that the set of quaternions of the form $a + b\,\mathbf{i}$ are a copy of the complex numbers living inside of $\mathbb{H}$. This is because for $a + b\,\mathbf{i}$ and $c + d\,\mathbf{i}$ we have that

$$(a + b\mathbf{i}) * (c + d\mathbf{i}) = (ac - bd) + (ad + bc)\mathbf{i} \overset{1\text{-}1}{\leftrightsquigarrow} (a + b\,i) \cdot (c + d\,i) = (ac - bd) + (ad + bc)i$$

whether we think of them as quaternions or complex numbers.[V] Moreover, for such elements both meanings of the conjugate and addition also coincide. Therefore, we do not feel the least bit guilty of saying that $\mathbb{C}$ *is* a subset of $\mathbb{H}$ given by the set of quaternions of the form $a + b\mathbf{i}$ called *complex quaternions*.[VI]

We now make more explicit what we referred to earlier in Remark 3.71 on page 107.

DEFINITION 3.84. The length of a quaternion q, denoted $\|\,\mathsf{q}\,\|$, is the length of the corresponding element of $\mathbb{R}^4$.

[V] In fact the only thing that we used here was that $i^2 = -1$.

[VI] The only possible cause for confusion is that the pure part of the quaternion $a + b\mathbf{i}$ is $b\mathbf{i}$ (it includes the vector $\mathbf{i}$!) while the imaginary part of the complex number $a + bi$ is just the number b.

That is if

$$\mathsf{q} = a + \mathbf{v} = a + v_1\,\mathbf{i} + v_2\,\mathbf{j} + v_3\,\mathbf{k} \Rightarrow \|\,\mathsf{q}\,\|^2 = a^2 + \|\,\mathbf{v}\,\|^2 = a^2 + v_1^2 + v_2^2 + v_3^2$$

where $\|\,\mathbf{v}\,\|$ has its usual meaning as the length of an element of $\mathbb{R}^3$. Like the length in every dimension, we have for all real scalars c

$$\|\,c \cdot \mathsf{q}\,\| = |\,c\,| \cdot \|\,\mathsf{q}\,\| \tag{20}$$

where $|\,c\,|$ is the absolute value of the real number c. In particular, if $c \geq 0$, we have that $\|\,c \cdot \mathsf{q}\,\| = c\,\|\,\mathsf{q}\,\|$.

Definition 3.85. The set of *unit quaternions* or the *3-sphere* is the set of quaternions of length 1. This is denoted as $\mathbb{S}^3$.[VII] The subset of $\mathbb{S}^3$ of pure quaternions is (identified with) $\mathbb{S}^2$, the two-sphere, and of complex quaternions is denoted by $\mathbb{S}^1$, the circle.

In symbols we therefore have

$$\begin{aligned}\mathbb{S}^3 &= \{a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k} : a^2 + b^2 + c^2 + d^2 = 1\},\\ \mathbb{S}^2 &= \{b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k} : b^2 + c^2 + d^2 = 1\},\ \text{and}\\ \mathbb{S}^1 &= \{a + b\,\mathbf{i} : a^2 + b^2 = 1\}.\end{aligned}$$

As in any dimension we can project any non-zero quaternion to $\mathbb{S}^3$ by dividing the given quaternion by its length. This is because

$$\left\|\,\frac{\mathsf{q}}{\|\,\mathsf{q}\,\|}\,\right\| = \frac{1}{\|\,\mathsf{q}\,\|}\,\|\,\mathsf{q}\,\| = 1.$$

Here we used (20) to pull out the positive scalar $\frac{1}{\|\,\mathsf{q}\,\|}$. In $\mathbb{R}^3$ and $\mathbb{R}^2$ you could picture this as a retraction of a point along the line connecting that point to the origin to the point where that line first meets the sphere. It is known as the *radial projection*. Like in 3-dimensions, the radial projection of q in four dimensions is denoted $\widehat{\mathsf{q}}$. Therefore, in summary,

$$\widehat{\mathsf{q}} = \frac{\mathsf{q}}{\|\,\mathsf{q}\,\|}.$$

Example 3.86. Let $\mathsf{q} = 1 + 3\,\mathbf{i} + 3\,\mathbf{j} - 9\,\mathbf{k}$, then

$$\|\,\mathsf{q}\,\|^2 = (1)^2 + (3)^2 + (3)^2 + (-9)^2 = 1 + 9 + 9 + 81 = 100$$

and so $\|\,\mathsf{q}\,\| = \sqrt{100} = 10$. Therefore the radial projection of q onto $\mathbb{S}^3$ is

$$\widehat{\mathsf{q}} = \frac{1}{10}\,(1 + 3\,\mathbf{i} + 3\,\mathbf{j} - 9\,\mathbf{k})\,.$$

Exercise 3.33. Show that

$$\begin{aligned}\pi :\, &\mathbb{R}^4 \setminus \{\mathbf{0}\} \to \mathbb{R}^4 \setminus \{\mathbf{0}\}\\ :\, &\mathsf{q} \mapsto \widehat{\mathsf{q}}\end{aligned}$$

is a projection whose image is $\mathbb{S}^3$.

SOLUTION

Exercise 3.34. Make a diagram which will help you keep track of which sets (from $\mathbb{S}^3$, $\mathbb{S}^2$, $\mathbb{S}^1$, $\mathbb{R}^3$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{C}^\times$, $\mathbb{H}^\times$) are contained inside of each other.

SOLUTION

Exercise 3.35. Calculate the length of each quaternion. Which are elements of $\mathbb{S}^3$? Which are elements of $\mathbb{S}^2$? Which are elements of $\mathbb{S}^1$? Which are real (i.e. elements of $\mathbb{R}$)? Which are pure (i.e. elements of $\mathbb{R}^3$)? (The number ϕ below is the golden ratio defined in Equation (10) back on page 99.)

SOLUTION

(1) 5:

(2) $\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\mathbf{i}$

(3) $\frac{1}{2} - \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}$

(4) $\frac{3}{7}\mathbf{i} + \frac{2}{7}\mathbf{j} - \frac{6}{7}\mathbf{k}$

(5) $\mathbf{i}$

(6) $\frac{\sqrt{2}}{2} - \frac{\sqrt{2}\mathbf{j}}{2}\mathbf{j}$

(7) $\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{j}$

(8) [VIII] $\frac{\phi}{2} + \frac{(\phi-1)}{2}\mathbf{i} + \frac{1}{2}\mathbf{j}$

(9) -1

(10) $\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\mathbf{i} + \frac{\sqrt{2}}{2}\mathbf{j} - \frac{\sqrt{2}}{2}\mathbf{k}$

Remark 3.87. We have already introduced several subsets of $\mathbb{H}$. Let's organize all these here for easy access. For a quaternion

$$\mathsf{q} = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$$

we have that q is in

$$\mathbb{R} \Leftrightarrow b = c = d = 0 \qquad \mathbb{R}^3 \Leftrightarrow a = 0$$
$$\mathbb{C} \Leftrightarrow c = d = 0 \qquad \mathbb{S}^3 \Leftrightarrow a^2 + b^2 + c^2 + d^2 = 1$$
$$\mathbb{S}^2 \Leftrightarrow a = 0, b^2 + c^2 + d^2 = 1 \qquad \mathbb{S}^1 \Leftrightarrow c = 0, d = 0, a^2 + b^2 = 1.$$

Of course we have

$$\mathbb{R} \cap \mathbb{R}^3 = \{0\} \qquad \mathbb{R}, \mathbb{S}^1 \subseteq \mathbb{C}$$
$$\mathbb{S}^1, \mathbb{S}^2 \subseteq \mathbb{S}^3 \qquad \mathbb{S}^1 = \mathbb{C} \cap \mathbb{S}^3$$
$$\mathbb{S}^1 \cap \mathbb{S}^2 = \{\pm\mathbf{i}\} \qquad \mathbb{S}^2 = \mathbb{R}^3 \cap \mathbb{S}^3.$$

Moreover, all of these sets are closed under conjugation, i.e. $\overline{\mathsf{q}}$ is in the set if and only if q is.

[VIII] You should do this one without using an actual value for ϕ but only using the fact that $\phi^2 = \phi + 1$.

Of course the length of a quaternion is a real number greater than or equal to 0, which equals zero only for the zero quaternion, $\mathsf{q} = 0$.

PROPOSITION 3.88. *For any quaternions* q *and* p *we have*

(1) $\mathsf{q} * \overline{\mathsf{q}} = \overline{\mathsf{q}} * \mathsf{q} = \| \mathsf{q} \|^2$

(2) $\overline{\mathsf{q} * \mathsf{p}} = \overline{\mathsf{p}} * \overline{\mathsf{q}}$.

PROOF.

(1) Write $\mathsf{q} = q_R + \mathbf{q}_P$ as the sum of its real and pure parts. Of course

$$\| \mathsf{q} \|^2 = q_R^2 + \| \mathbf{q}_P \|^2 ,$$

so we are trying to show that $\mathsf{q} * \overline{\mathsf{q}} = q_R^2 + \| \mathbf{q}_P \|^2$ (we'll talk about the other half of this in a minute). But using the formulas,

$$\begin{aligned}
\mathsf{q} * \overline{\mathsf{q}} &= (q_R + \mathbf{q}_P) * (q_R - \mathbf{q}_P) \\
&= \left(q_R^2 - \mathbf{q}_P \bullet (-\mathbf{q}_P)\right) + \left(q_R(-\mathbf{q}_P) + q_R\mathbf{q}_P + \cancel{\mathbf{q}_P \times \mathbf{q}_P}\right) && \text{(Vector cross itself = } \mathbf{0}\text{)} \\
&= \left(q_R^2 + \mathbf{q}_P \bullet \mathbf{q}_P\right) + \left(\cancel{-q_R\mathbf{q}_P} + \cancel{q_R\mathbf{q}_P}\right) \\
&= q_R^2 + \| \mathbf{q}_P \|^2 && \text{(Vector dot itself = length squared)}
\end{aligned}$$

as desired. To see that $\overline{\mathsf{q}} * \mathsf{q}$ also equals $\| \mathsf{q} \|^2$, we could do a similar calculation, or we could be slick and simply use the fact that both q and its conjugate, $\overline{\mathsf{q}}$ have the same length.

(2) To prove this second part, we note that for $\mathsf{q} = q_R + \mathbf{q_P}$ and $\mathsf{p} = p_R + \mathbf{p_P}$

$$\begin{aligned}
\overline{\mathsf{q} * \mathsf{p}} &= \overline{(q_R p_R - \mathbf{q}_P \bullet \mathbf{p}_P) + (q_R\mathbf{p}_P + p_R\mathbf{q}_P + \mathbf{q}_P \times \mathbf{p}_P)} \\
&= (q_R p_R - \mathbf{q}_P \bullet \mathbf{p}_P) - (q_R\mathbf{p}_P + p_R\mathbf{q}_P + \mathbf{q}_P \times \mathbf{p}_P) \\
&= (q_R p_R - \mathbf{q}_P \bullet \mathbf{p}_P) + (q_R(-\mathbf{p}_P) + p_R(-\mathbf{q}_P) + \mathbf{p}_P \times \mathbf{q}_P). \quad \text{(Using } \mathbf{q}_P \times \mathbf{p}_P = -\mathbf{p}_P \times \mathbf{q}_P\text{)} \qquad (21)
\end{aligned}$$

At the same time we have

$$\begin{aligned}
\overline{\mathsf{p}} * \overline{\mathsf{q}} &= (p_R - \mathbf{p}_P) * (q_R - \mathbf{q}_P) \\
&= (q_R\, p_R - \mathbf{q}_P \bullet \mathbf{p}_P) + (q_R(-\mathbf{p}_P) + p_R(-\mathbf{q}_P) + \mathbf{p}_P \times \mathbf{q}_P). \qquad (22)
\end{aligned}$$

Comparing (21) and (22) they are indeed equal and so what we started with, $\overline{\mathsf{q} * \mathsf{p}}$ and $\overline{\mathsf{p}} * \overline{\mathsf{q}}$ are equal, as desired. ■

In particular, since a pure quaternion is equal to negative its own conjugate, we have for a pure quaternions $\mathbf{v}$, that $\mathbf{v}^2 = -\mathbf{v} * \overline{\mathbf{v}} = -\| \mathbf{v} \|^2$. For instance, we have

$$(3\mathbf{i} + 5\mathbf{j} - 6\mathbf{k})^2 = -(3^2 + 5^2 + 6^2) = -(9 + 25 + 36) = -70.$$

Translating the remark right before the previous theorem, the first part of this theorem says $\mathsf{q} * \overline{\mathsf{q}}$ is a real number which is greater than or equal to 0. Moreover, this number is equal to 0 iff $\mathsf{q} = 0 (= 0 + 0\mathbf{i} + 0\mathbf{j} + 0\mathbf{k})$. As such, the following definition makes sense and it corresponds to the usual meaning of a length of a (vector) in $\mathbb{R}^4$.

THEOREM 3.89. *For any quaternion* $\mathsf{q} \neq 0$ *we have that*

$$\mathsf{q} * \left(\frac{\overline{\mathsf{q}}}{\| \mathsf{q} \|^2}\right) = 1 = \left(\frac{\overline{\mathsf{q}}}{\| \mathsf{q} \|^2}\right) * \mathsf{q}.$$

Proof. The assumption, $q \neq 0$, guarantees us that $\| q \|^2 \neq 0$. Therefore, we have that

$$q * \left(\frac{\overline{q}}{\| q \|^2} \right) = \frac{1}{\| q \|^2} \cdot (q * \overline{q}) = \frac{1}{\| q \|^2} \cdot \| q \|^2 = 1.$$

The proof that $\left(\frac{\overline{q}}{\| q \|^2} \right) * q$ also equals 1 is the same. ∎

As pointed out in the proof, the reason why we needed to require that $q \neq 0$ is to guarantee that we weren't dividing by 0.

Definition 3.90. For any $q \neq 0$, we define q^{-1} as the quantity $\frac{\overline{q}}{\| q \|^2}$ from the theorem.

If you want a general formula for the inverse of a quaternion

$$q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$$

it is simply

$$\begin{aligned} q^{-1} &= \frac{a}{a^2 + b^2 + c^2 + d^2} - \frac{b}{a^2 + b^2 + c^2 + d^2}\mathbf{i} - \frac{c}{a^2 + b^2 + c^2 + d^2}\mathbf{j} - \frac{d}{a^2 + b^2 + c^2 + d^2}\mathbf{k} \\ &= \frac{a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}}{a^2 + b^2 + c^2 + d^2}.\text{[IX]} \end{aligned}$$

Example 3.91. The quaternion $1 + \mathbf{i} + \mathbf{j}$ has inverse

$$q^{-1} = \frac{1}{3} - \frac{1}{3}\mathbf{i} - \frac{1}{3}\mathbf{j}.$$

Example /Theorem 3.92. $q^{-1} = \overline{q} \Leftrightarrow q \in \mathbb{S}^3$.

Proof.

⇒ Suppose that $q^{-1} = \overline{q}$. Then $q * \overline{q} = 1$. But we know that for any quaternion q, $q * \overline{q} = \| q \|^2$. Comparing equations shows that $\| q \|^2 = 1$. Since the length of a quaternion is always positive, this proves that $\| q \| = 1$.

⇐ Simply reverse the steps of the previous part. ∎

For the standard quaternions, **i**, **j**, and **k**, we have that their inverse is their negative. The following exercise shows that this classifies unit pure quaternions, i.e. elements of $\mathbb{S}^2$.

Exercise 3.36. Prove that a quaternion q is in $\mathbb{S}^2$ iff $q^{-1} = -q$.

Solution

[IX] We have never defined and never will define a quaternion such as $\frac{1}{1+\mathbf{i}+\mathbf{j}}$. You should always keep your **i**, **j**, and **k**'s in the numerator! In the case of quaternions you won't get yourself into too much trouble, but in group theory this is a bad habit you should get out of. The only time you really should still be writing one over something is if that something is a real number!

Exercise 3.37. A set X is *closed under multiplication* if for all $\mathsf{q}_1, \mathsf{q}_2 \in X$ we have that $\mathsf{q}_1 * \mathsf{q}_2$ is still in X. We say that X is closed under inversion if for all $a \in X$ we have that $\mathsf{q}^{-1} \in X$. Obviously $\mathbb{R}$ and $\mathbb{C}$ are closed under multiplication ($\mathbb{R}, \mathbb{C} \subseteq \mathbb{H}$.) and inversion.

Solution

(1) Show that $\mathbb{R}^3$ and $\mathbb{S}^2$ are closed under inversion.

(2) Show that $\mathbb{R}^3$ and $\mathbb{S}^2$ are **not** closed under multiplication. (You can do this in one shot by finding a quaternion which is in both sets but whose square is in neither. **i** know a really easy instance of this.)

(3) Show that $\mathbb{S}^1$ is closed under multiplication and inversion.

(4) We will see in Corollary 3.101 that $\mathbb{S}^3$ is closed under multiplication and inversion. Verify that it's closed under multiplication by taking four concrete explicit examples (don't use any theorems, just multiply them out and verify that their product still has length 1).

We have proven that q^{-1} isn't a sucky name since it is indeed a multiplicative inverse for q. Notice also that for unit quaternions q, that $\mathsf{q}^{-1} = \overline{\mathsf{q}}$. This is because the $\| \mathsf{q} \|^2$ term in the denominator is equal to 1 for unit quaternions. In particular, for unit pure quaternions (i.e. elements of $\mathbb{S}^2$) we have that $\mathbf{u}^{-1} = -\mathbf{u}$. As such, for such quaternions we have that

$$\begin{aligned} \mathbf{u}^2 &= \mathbf{u} * \mathbf{u} \\ &= \mathbf{u} * (-\mathbf{u}^{-1}) \\ &= -\mathbf{u} * \mathbf{u}^{-1} \\ &= -1. \end{aligned}$$

This generalizes the fact that $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2$ all equal -1 to some crazy unit pure quaternions such as

$$\frac{3}{7}\mathbf{i} + \frac{2}{7}\mathbf{j} - \frac{6}{7}\mathbf{k}.$$

Exercise 3.38. Verify "by hand" that

Solution

$$\left(\frac{3}{7}\mathbf{i} + \frac{2}{7}\mathbf{j} - \frac{6}{7}\mathbf{k}\right)^2 = -1.$$

More generally for all pure quaternions $\mathbf{v}$ we have that

$$\mathbf{v}^2 = -\| \mathbf{v} \|^2 .$$

Of course we could see these identities in terms of the original definitions of the quaternion multiplication.

Example/Theorem 3.93. Two non-zero quaternions, p and q anti-commute if and only if they are orthogonal pure quaternions.

Proof. Denote

$$\mathsf{p} = p_R + \mathbf{p}_P \quad \text{and} \quad \mathsf{q} = q_R + \mathbf{q}_P.$$

If p and q also commute, then

$$\begin{aligned} \mathsf{p} * \mathsf{q} &= \mathsf{q} * \mathsf{p} \\ &= -\mathsf{p} * \mathsf{q} \\ &\Rightarrow \mathsf{p} * \mathsf{q} = 0 \\ &\Rightarrow \mathsf{p} = 0 \text{ or } \mathsf{q} = 0 \end{aligned}$$

which contradicts our assumption that p and q are non-zero. Thus $\mathbf{p}_P$ and $\mathbf{q}_P$ cannot be parallel to each other. A direct calculation shows us that

$$\mathsf{q} * \mathsf{p} = -\mathsf{p} * \mathsf{q} \Rightarrow 0 = \mathsf{q} * \mathsf{p} + \mathsf{p} * \mathsf{q} = 2\Big((q_R p_R - \mathbf{q}_P \bullet \mathbf{p}_P) + (q_R \mathbf{p}_P + p_R \mathbf{q}_P)\Big).$$

Therefore the pure and real parts of the above must be zero. Let's first see what setting the pure part equal to **0** tells us:

$$q_R \mathbf{p}_P + p_R \mathbf{q}_P = \mathbf{0}.$$

Since $\mathbf{q}_P$ and $\mathbf{p}_P$ are not parallel, we must have that both p_R and q_R are zero. That is both p and q must have real parts equal to 0 and, as such, are pure. Now, glancing at the real part of the above, we see that

$$\cancel{q_R p_R} - \mathbf{q}_P \bullet \mathbf{p}_P = 0 \Rightarrow \mathbf{q}_P \bullet \mathbf{p}_P = 0 \Rightarrow \mathbf{q}_P \perp \mathbf{p}_P.$$

In conclusion, two quaternions anti-commute iff they are orthogonal pure quaternions. ■

In summary, for p and q quaternions we have that

- p and q commute iff their pure parts are parallel.[x]
- p and q anti-commute iff they are orthogonal and pure.

COROLLARY 3.94. *If* $\mathsf{q} \in \mathbb{S}^3$, $\mathbf{u} \in \mathbb{S}^2$, *and* $\mathbf{Pu}\,\mathsf{q} \perp \mathbf{u}$, *then*

(1) $\mathbf{u}\mathsf{q}\mathbf{u}^{-1} = \mathsf{q}^{-1}$, *and*

(2) $(\mathbf{u}\mathsf{q})^2 = -1 = (\mathsf{q}\mathbf{u})^2$.

PROOF.

(1) By Lemma 3.93 we know that $\mathbf{Pu}\,\mathsf{q}$ and **u** anti-commute. This means that

$$\begin{aligned} \mathbf{u}\mathsf{q} &= \mathbf{u}\,(\mathbf{Re}\,\mathsf{q} + \mathbf{Pu}\,\mathsf{q}) \\ &= \mathbf{u}\,\mathbf{Re}\,\mathsf{q} + \mathbf{u}\,\mathbf{Pu}\,\mathsf{q} && \text{(Distrib)} \\ &= \mathbf{Re}\,\mathsf{q}\mathbf{u} + \mathbf{u}\,\mathbf{Pu}\,\mathsf{q} && (\mathbf{u} \text{ commutes w/real } \mathbf{Re}\,\mathsf{q}) \\ &= \mathbf{Re}\,\mathsf{q}\mathbf{u} - \mathbf{Pu}\,\mathsf{q}\mathbf{u} && (\mathbf{u} \text{ and } \mathbf{Pu}\,\mathsf{q} \text{ anti-commute}) \\ &= (\mathbf{Re}\,\mathsf{q} - \mathbf{Pu}\,\mathsf{q})\,\mathbf{u} && \text{(Distrib)} \\ &= \bar{\mathsf{q}}\mathbf{u}. \end{aligned}$$

i.e.

$$\mathbf{u}\mathsf{q} = \bar{\mathsf{q}}\mathbf{u}. \tag{23}$$

[x] We consider every vector to be parallel to the zero vector. This goes with the fact that every quaternion commutes with every real quaternion.

But since $q \in \mathbb{S}^3$ we also know that $\overline{q} = q^{-1}$. Thus (23) becomes

$$\mathbf{u}q = q^{-1}\mathbf{u}. \tag{24}$$

Since $\mathbf{u} \in \mathbb{S}^2$, it is, in particular, not equal to 0. Therefore $\mathbf{u}^{-1}$ is defined (in fact it's equal to $-\mathbf{u}$), and we have

$$\mathbf{u}q\mathbf{u}^{-1} = q^{-1} \tag{25}$$

as desired.

(2) Since

$$\mathbf{u} \in \mathbb{S}^2 \Rightarrow \mathbf{u} \in \mathbb{S}^3 \cap \mathbb{R}^3 \Rightarrow \mathbf{u}^{-1} = \overline{\mathbf{u}}, \overline{\mathbf{u}} = -\mathbf{u} \Rightarrow \mathbf{u}^{-1} = -\mathbf{u}$$

we can rearrange equation (25) to say that

$$\begin{aligned} \mathbf{u}q\mathbf{u}^{-1} &= q^{-1} \\ \Rightarrow -\mathbf{u}q\mathbf{u} &= q^{-1} \\ \Rightarrow -\mathbf{u}q\mathbf{u}q = 1 \quad &\text{and} \quad -q\mathbf{u}q\mathbf{u} = 1 \\ \Rightarrow -(\mathbf{u}q)^2 = 1 \quad &\text{and} \quad -(q\mathbf{u})^2 = 1 \\ \Rightarrow (\mathbf{u}q)^2 = -1 \quad &\text{and} \quad (q\mathbf{u})^2 = -1. \end{aligned}$$

as desired. ∎

THEOREM 3.95. **($\mathbb{C} \times \mathbb{C}$ Version)** *For quaternions* (z, w) *and* $(z_1, w_1) \in \mathbb{C} \times \mathbb{C}$ *we have*

$$(z, w) * (z_1, w_1) = (zz_1 - w\overline{w}_1, zw_1 + w\overline{z}_1). \qquad ∎$$

EXERCISE 3.39. Prove the theorem, 3.95

SOLUTION

We will only use this formula only rarely. But its use is to prove possibly the most important fact about quaternions.

EXAMPLE 3.96.

$$\begin{aligned} (1 + i, 2 + i) * (i, 1 - i) &= ((1 + i)i - (2 + i)(1 + i), (1 + i)(1 - i) + (2 + i)(-i)) \\ &= ((i - 1) - (1 + 3i), 2 - 2i + 1) \\ &= (-2 - 2i, 3 - 2i). \end{aligned}$$

This corresponds to the fact that

$$(1 + \mathbf{i} + 2\mathbf{j} + \mathbf{k}) * (\mathbf{i} + \mathbf{j} - \mathbf{k}) = -2 - 2\mathbf{i} + 3\mathbf{j} - 2\mathbf{k}.$$

Exercise 3.40. Verify that in this situation that we have for complex numbers z and w that

$$(z,0)*(w,0)=(z\cdot w,0) \quad \text{and} \quad (0,z)*(0,w)=(-z\cdot w,0).$$

SOLUTION

Exercise 3.41. Translate the quaternions

(1) $(5,\mathbf{i}+2\mathbf{j})$

(2) $(1,\mathbf{i}+\mathbf{k})$

(3) $3+2i+5j+k$

(4) $1+i$

(5) $(3+i,2+i)$

(6) $(5+i,6)$

SOLUTION

into its other versions. For instance, the quaternion $(1,2\mathbf{i}+3\mathbf{j}+4\mathbf{k})$ is in $\mathbb{R}\times\mathbb{R}^3$ notation. This same quaternion in $\mathbb{R}^4$ notation is $1+2i+3j+4k$ and in $\mathbb{C}\times\mathbb{C}$ notation it is $(1+2i,3+4i)$. If you want to make this really easy to grade make three columns corresponding to the three versions and write the original in, say, pen. Then write the other two versions in, say, pencil. Therefore, your answer for the same I just gave you would look like:

	$\mathbb{R}^4$	$\mathbb{R}\times\mathbb{R}^3$	$\mathbb{C}\times\mathbb{C}$
1.	$1+2i+3j+4k$	$(1,2\mathbf{i}+3\mathbf{j}+4\mathbf{k})$	$(1+2i,3+4i)$

Exercise 3.42. Let

$$q_1=(2,3\mathbf{i}+\mathbf{k}), \quad q_2=5+j+2k, \quad \text{and} \quad q_3=(1+i,3+i).$$

Calculate both q_1*q_2 and q_2*q_3 in each of the three ways (so you need to do a total of 6 calculations) and verify for each that you get the same answer.

SOLUTION

Remark 3.97. Like everything else, it's tempting to pick "favorites." What is your favorite version of the quaternions? While tempting this really should be avoided! The ideal thought process is

- What does each version bring to the table?
- What are the strengths and weakness of each version?

Just like I pointed out in the introduction, even though I'm sure if you ask a carpenter what their favorite tool is they would have an opinion, they certainly will not try to use that tool for a job it is not meant for. I love my power screwdriver, but I will not try to put a nail into a wall or cut a string with it!

Theorem 3.98. *The multiplication of quaternions is associative. That is for all quaternions q_1, q_2, and q_3 we have that*

$$(q_1*q_2)*q_3=q_1*(q_2*q_3).$$

Proof. Write $q_1 = (z, w)$, $q_2 = (z_1, w_1)$ and $q_3 = (z_2, w_2)$ (both of which are in $\mathbb{C} \times \mathbb{C}$ as above). Then (rev up for a long calculation!!) $(q_1 * q_2) * q_3$ is

$$\begin{aligned}
&= \left(zz_1 - w\overline{w}_1, zw_1 + w\overline{z}_1\right) * q_3 \\
&= \left((zz_1 - w\overline{w}_1)z_2 - (zw_1 + w\overline{z}_1)\overline{w}_2, (zz_1 - w\overline{w}_1)w_2 + (zw_1 + w\overline{z}_1)\overline{z}_2\right) \\
&= \left(zz_1z_2 - w\overline{w}_1z_2 - zw_1\overline{w}_2 - w\overline{z}_1\overline{w}_2, zz_1w_2 - w\overline{w}_1w'' + zw_1\overline{z}_2' + w\overline{z}_1\overline{z}_2\right).
\end{aligned}$$

and at the same time we have that $q_1 * (q_2 * q_3)$ is

$$\begin{aligned}
&= q_1 * (z_1z_2 - w_1\overline{w}_2, z_1w_2 + w_1\overline{z}_2) \\
&= \left(z(z_1z_2 - w_1\overline{w}_2) - w\overline{(z_1w_2 + w_1\overline{z}_2)}, z(z_1w_2 + w_1\overline{z}_2) + w\overline{(z_1z_2 - w_1\overline{w}_2)}\right) \\
&= \left(zz_1z_2 - zw_1\overline{w}_2 - w\overline{z}_1\overline{w}_2 - w\overline{w}_1z_2, zz_1w_2 + zw_1\overline{z}_2 + w\overline{z}_1\overline{z}_2 - w\overline{w}_1w_2\right).
\end{aligned}$$

Comparing the two equations gives the desired equality. ∎

Corollary 3.99. *For any two quaternions,* q *and* p *we have*

$$\|q * p\| = \|q\| \cdot \|p\| \quad \textit{and} \quad \|q * p\|^2 = \|q\|^2 \cdot \|p\|^2 .^{\text{XI}}$$

Proof. In Proposition 3.88 we saw that

$$\|q\|^2 = q * \overline{q} \text{ and } \overline{q * p} = \overline{p} * \overline{q}.$$

Therefore

$$\begin{aligned}
\|q * p\|^2 &= (q * p) * (\overline{q * p}) \\
&= (q * p) * (\overline{p} * \overline{q}) && (\overline{q * p} = \overline{p} * \overline{q}) \\
&= q * (p * \overline{p}) * \overline{q} && \text{(Associativity)} \\
&= q * \|p\|^2 * \overline{q} && (p * \overline{p} = \|p\|^2) \\
&= \|p\|^2 * (q * \overline{q}) && (\|p\|^2 \in \mathbb{R}) \\
&= \|p\|^2 \cdot \|q\|^2 .
\end{aligned}$$

We can switch the order of this last multiplication because in it we are multiplying together two real numbers. This proves the second formula. Taking square roots proves the first. ∎

Example 3.100. We verify the corollary, Corollary 3.99, for a concrete example. Let's take

$$q = 1 + \mathbf{i} + \mathbf{j} \quad \text{and} \quad p = 1 + 2\,\mathbf{i} + \mathbf{k}$$

as we did back in Example 3.80. Back in that example we saw that

$$q * p = -1 + 4\,\mathbf{i} - \mathbf{k} \Rightarrow \|q * p\|^2 = 1^2 + 4^2 + 1^2 = \boxed{18}.$$

[XI] On the left hand side of each equation, we are multiplying two quaternions and taking their length, and on the right hand side of each equation we are performing the old fashioned multiplication of two numbers.

Meanwhile we get the same result if we calculate this the other way:

$$\|q\|^2 = 1+1+1 = 3 \quad \text{and} \quad \|p\|^2 = 1+4+1 = 6 \Rightarrow \|q\|^2 \cdot \|p\|^2 = 3 \cdot 6 = \boxed{18},$$

and the corollary is verified.

The following Corollary is very important. It says for any $q, p \in \mathbb{S}^3$ and $q * p$ and p are also in $\mathbb{S}^3$. (I have already asked you to verify some cases of this in Exercise 3.37.) Later, in Chapter 6, it will be translated as saying that $\mathbb{S}^3$ is a subgroup of $\mathbb{H}^\times$.

COROLLARY 3.101. *$\mathbb{S}^3$ is closed under*

(1) *under multiplication and*

(2) *inversion.*

PROOF.

(1) Let q and $p \in \mathbb{S}^3$. Then, by definition, $\|q\| = \|p\| = 1$. We need to show that $\|q * p\| = 1$. But using Corollary 3.99 this is easy:

$$\|p * q\| = 1 \cdot 1 = 1.$$

(2) This is left for you in Exercise 3.43. ■

EXERCISE 3.43. Use Corollary 3.99 to prove that for any quaternion q we have that

$$\|q^{-1}\| = \frac{1}{\|q\|}.$$

(This formula is not too difficult for you to prove directly, but I would prefer it if you could prove it by only using the Corollary and the fact that $\|1\| = 1$.) Conclude that $\mathbb{S}^3$ is closed under inversion (that is if $q \in \mathbb{S}^3$, then $q^{-1} \in \mathbb{S}^3$ is as well).

SOLUTION

Associativity also allows us to easily prove many identities and formulas about vector cross products and dot products. For instance, let's prove some interesting results. Of course since the cross and dot products have explicit formulas these can be derived after a grueling by-hand calculation. (We won't really use these results but they are pretty cool since we have managed to derive them in a purely algebraic way)

COROLLARY 3.102.

(1) $\mathbf{v} \times \mathbf{w} = \mathbf{0} \Leftrightarrow \mathbf{v} \asymp \mathbf{w}$.

(2) *For any three vectors* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *we have the following triple product rule:*

$$\mathbf{v} \times (\mathbf{u} \times \mathbf{w}) = (\mathbf{w} \bullet \mathbf{v})\mathbf{u} - (\mathbf{v} \bullet \mathbf{u})\mathbf{w}.$$

PROOF.

(1) If $\mathbf{v} \times \mathbf{w} = \mathbf{0}$, then we have that

$$\mathbf{v} * \mathbf{w} = \mathbf{w} * \mathbf{v} = -\mathbf{v} \bullet \mathbf{w}.$$

As such we have

$$(\mathbf{v} * \mathbf{w}) * \mathbf{v} = -(\mathbf{v} \bullet \mathbf{w})\mathbf{v} \quad \text{and} \quad (\mathbf{w} * \mathbf{v}) * \mathbf{v} = \mathbf{w} * (\mathbf{v} * \mathbf{v}) = -\|\mathbf{v}\|^2 \mathbf{w}$$

must be equal. As such

$$\mathbf{w} = \frac{(\mathbf{v} \bullet \mathbf{w})}{\|\mathbf{v}\|^2}\mathbf{v}$$

and $\mathbf{v}$ and $\mathbf{w}$ are proportional.

(2) In this proof we will make extensive use of the identities

$$2\mathbf{v} \times \mathbf{w} = (a + \mathbf{v}) * (b + \mathbf{w}) - (b + \mathbf{w}) * (a + \mathbf{v}) \tag{26}$$

$$-2\mathbf{v} \bullet \mathbf{w} = \mathbf{v} * \mathbf{w} + \mathbf{w} * \mathbf{v} \tag{27}$$

which we have used several times already (their proofs are via a direct calculation). Then for pure quaternions $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{u}$ we have that

$$\begin{aligned}
(\mathbf{u} * \mathbf{v}) * \mathbf{w} &= (-\mathbf{v} * \mathbf{u} - 2\mathbf{v} \bullet \mathbf{u}) * \mathbf{w} && \text{(Identity (27) Above)} \\
&= (-\mathbf{v} * \mathbf{u}) * \mathbf{w} - 2(\mathbf{v} \bullet \mathbf{u})\mathbf{w} && \text{(Distributive Property)} \\
&= -\mathbf{v} * (\mathbf{u} * \mathbf{w}) - 2(\mathbf{v} \bullet \mathbf{u})\mathbf{w}. && \text{(Associative Property)}
\end{aligned} \tag{28}$$

Meanwhile the original quantity $(\mathbf{u} * \mathbf{v}) * \mathbf{w}$ is equal to $\mathbf{u} * (\mathbf{v} * \mathbf{w})$ by associativity. Now working with this quantity we see that

$$\begin{aligned}
\mathbf{u} * (\mathbf{v} * \mathbf{w}) &= \mathbf{u} * (-\mathbf{w} * \mathbf{v} - 2\mathbf{w} \bullet \mathbf{v}) && \text{(Identity (27) Above)} \\
&= -\mathbf{u} * (\mathbf{w} * \mathbf{v}) - 2(\mathbf{w} \bullet \mathbf{v})\mathbf{u} && \text{(Distributive Property)} \\
&= -(\mathbf{u} * \mathbf{w}) * \mathbf{v} - 2(\mathbf{w} \bullet \mathbf{v})\mathbf{u}. && \text{(Associative Property)}
\end{aligned} \tag{29}$$

Now we subtract these two equations, (28) from (29), we see that

$$\begin{aligned}
0 &= \Big(\mathbf{v} * (\mathbf{u} * \mathbf{w}) + 2(\mathbf{v} \bullet \mathbf{u})\mathbf{w}\Big) - \Big((\mathbf{u} * \mathbf{w}) * \mathbf{v} + 2(\mathbf{w} \bullet \mathbf{v})\mathbf{u}\Big) \\
&= \mathbf{v} * (\mathbf{u} * \mathbf{w}) - (\mathbf{u} * \mathbf{w}) * \mathbf{v} + 2(\mathbf{v} \bullet \mathbf{u})\mathbf{w} - 2(\mathbf{w} \bullet \mathbf{v})\mathbf{u}.
\end{aligned} \tag{30}$$

Now the quantity in blue is of the form of Identity (26) above (with

$$\mathbf{u} * \mathbf{w} = -\mathbf{u} \bullet \mathbf{w} + \mathbf{u} \times \mathbf{w}$$

being one of the two quaternions in the identity and $\mathbf{v} = \mathbf{0} + \mathbf{v}$ being the other). Therefore, the identity says that the quantity in blue is

$$2\mathbf{v} \times (\mathbf{u} \times \mathbf{w}).$$

Swapping this in in (30) we see that

$$0 = 2\mathbf{v} \times (\mathbf{u} \times \mathbf{w}) - 2(\mathbf{v} \bullet \mathbf{u})\mathbf{w} - 2(\mathbf{w} \bullet \mathbf{v})\mathbf{u} \Rightarrow \mathbf{v} \times (\mathbf{u} \times \mathbf{w}) = (\mathbf{w} \bullet \mathbf{v})\mathbf{u} - (\mathbf{v} \bullet \mathbf{u})\mathbf{w}$$

as desired. ■

THEOREM 3.103 (Spherical Coordinantes for $\mathbb{S}^3$). *For every element* $\mathsf{q} \in \mathbb{S}^3$ *there exists a* $\widehat{\mathbf{u}} \in \mathbb{S}^2$ *and* $0 \leq \theta \leq \pi$ *so that*

$$\mathsf{q} = \cos\theta + \sin\theta\,\widehat{\mathbf{u}}.$$

PROOF. Let $\mathsf{q} = q_R + \mathbf{q}_P \in \mathbb{S}^3$. If the pure part of q, $\mathbf{q}_P$, is zero then we'd have $q_R^2 = 1$ and so $q_R = \pm 1$ and we are done. Otherwise define $\widehat{\mathbf{u}}$ to be normalization of $\mathbf{q}_P$ given by

$$\widehat{\mathbf{u}} := \frac{\mathbf{q}_P}{\|\,\mathbf{q}_P\,\|}.$$

We know that

$$\|\,\mathsf{q}\,\|^2 = q_R^2 + \|\,\mathbf{q}_P\,\|^2 = 1.$$

As such there is an angle θ between 0 and 2π so that $\cos\theta = q_R$ and $\sin\theta = \|\,\mathbf{q}_P\,\|$ (the pair of real numbers $(q_R, \|\,\mathbf{q}_P\,\|)$ live on the unit circle). But since the latter number is positive we can take θ to be less than equal to π. ∎

This theorem says that every unit quaternion lives in some circle. Given such a unit quaternion, we can find θ and $\widehat{\mathbf{u}}$ by noting that

$$\theta = \cos^{-1}(\mathbf{Re}\,\mathsf{q}) \quad \text{and} \quad \widehat{\mathbf{u}} = \frac{1}{\sin\theta}\,\mathbf{Pu}\,\mathsf{q} \quad \text{or} \quad \widehat{\mathbf{u}} = \frac{1}{\sqrt{1-\mathbf{Re}\,\mathsf{q}^2}}\,\mathbf{Pu}\,\mathsf{q}.$$

Note also that except for the pesky quaternions, $+1$ and -1 [XII], these θ and $\widehat{\mathbf{u}}$ are unique.

EXAMPLE 3.104.

(1) Let q be the quaternion

$$\mathsf{q} = \frac{\sqrt{3}}{2} + \frac{3}{11}\mathbf{i} + \frac{1}{11}\mathbf{j} - \frac{9}{22}\mathbf{k}.$$

We find the θ and $\widehat{\mathbf{u}}$ which exist via Theorem 3.103. It's straight forward to show that $\mathsf{q} \in \mathbb{S}^3$ (see Exercise 3.48). Looking at the pure part, since

$$\frac{\sqrt{3}}{2} = \cos\frac{\pi}{6}$$

we see that $\theta = \frac{\pi}{6}$. Therefore $\widehat{\mathbf{u}}$ must be

$$\frac{1}{\sin\frac{\pi}{6}}\left(\frac{3}{11}\mathbf{i} + \frac{1}{11}\mathbf{j} - \frac{9}{22}\mathbf{k}\right) = \frac{6}{11}\mathbf{i} + \frac{2}{11}\mathbf{j} - \frac{9}{11}\mathbf{k}.$$

Indeed this dude really is in $\mathbb{S}^2$ since we have

$$\frac{1}{11^2}\left(6^2 + 2^2 + 9^2\right) = \frac{1}{121}(36 + 4 + 81) = 1.$$

(note that we knew this already!)

(2) The quaternion $\mathsf{q} = \mathbf{j}$ is of the form of the theorem by taking $\theta = \pi/2$ and $\mathbf{v} = \mathbf{j}$.

[XII] For ± 1 any $\widehat{\mathbf{u}}$ will due since for all $\widehat{\mathbf{u}}$ we have

$$1 = \cos 0 + \sin 0\,\widehat{\mathbf{u}} \quad \text{and} \quad -1 = \cos\pi + \sin\pi\,\widehat{\mathbf{u}}.$$

Let $\asymp$ be the relation on $\mathbb{R}^3 \setminus \{\mathbf{0}\}$ from Example 2.72.1 back on page 50 which related two non-zero vectors iff they were parallel. We also consider the zero vector, $\mathbf{0}$, to be be parallel to every vector. (This breaks the transitivity of $\asymp$.)

EXAMPLE 3.105. A quick example to make sure we are all on the same page.

$$3\mathbf{i} + \mathbf{j} \asymp 6\mathbf{i} + 2\mathbf{j} \asymp \frac{3}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} \asymp \mathbf{0} \asymp 2\mathbf{k} \asymp 9\mathbf{k}$$

but no vector before the zero vector (in red) is parallel to any vector after the zero vector (blue), in particular,

$$3\mathbf{i} + \mathbf{j} \not\asymp 9\mathbf{k}.$$

DEFINITION 3.106. Let $\mathbf{v} \in \mathbb{R}^3$ and $\widehat{\mathbf{u}} \in \mathbb{S}^2$ (so $\mathbf{v}$ is pure and $\widehat{\mathbf{u}}$ is pure unit), then define

$$\mathbb{C}_{\mathbf{v}} \stackrel{\text{def}}{=} \{\mathsf{q} : \mathbf{Pu}\,\mathsf{q} \asymp \mathbf{v}\} = \{\mathsf{q} : \mathsf{q} = a + b\mathbf{v} \text{ for some } a, b \in \mathbb{R}\},$$
$$\mathbb{C}_{\mathbf{v}}^{\times} \stackrel{\text{def}}{=} \mathbb{C}_{\mathbf{v}} \setminus \{0\}, \text{ and}$$
$$\mathbb{S}^1_{\widehat{\mathbf{u}}} \stackrel{\text{def}}{=} \mathbb{C}_{\widehat{\mathbf{u}}} \cap \mathbb{S}^3 = \{a + b\widehat{\mathbf{u}} : a^2 + b^2 = 1\} = \{\cos(\theta) + \sin(\theta)\,\widehat{\mathbf{u}} : 0 \le \theta < 2\pi\}.$$

Under these definitions, we have that $\mathbb{C}_i = \mathbb{C}$ and $\mathbb{S}^1_i = \mathbb{S}^1$. Moreover since $\mathbf{0}$ is parallel to every vector in $\mathbb{R}^3$ we have that $\mathbb{C}_{\mathbf{0}} = \mathbb{H}$. We've already talked about how $\mathbb{C}_i$ was a copy of the complex numbers, but moreover this is the case for $\mathbb{C}_{\widehat{\mathbf{u}}}$ whenever $\widehat{\mathbf{u}} \in \mathbb{S}^2$. In this way we can (almost) think of $\mathbb{H}$ as an $\mathbb{S}^2$'s worth of complex planes, and $\mathbb{S}^3$ as an $\mathbb{S}^2$'s worth of circles. [XIII]

Exercise 3.44. Throughout this problem $\mathbf{v}$ and $\mathbf{v}'$ are non-zero 3-vectors and $\widehat{\mathbf{u}}, \widehat{\mathbf{u}}' \in \mathbb{S}^2$. Prove that

(1) Every vector $\neq \mathbf{0}$ is parallel to exactly two anti-polar[XIV] vectors in $\mathbb{S}^2$;
(2) $\mathbb{C}_{\mathbf{v}} = \mathbb{C}_{\mathbf{v}'} \Leftrightarrow \mathbf{v} \asymp \mathbf{v}'$;
(3) $\mathbb{S}^1_{\widehat{\mathbf{u}}} = \mathbb{S}^1_{\widehat{\mathbf{u}}'} \Leftrightarrow \widehat{\mathbf{u}} \asymp \widehat{\mathbf{u}}'$;
(4) $\mathbf{v} \not\asymp \mathbf{v}' \Rightarrow \mathbb{C}_{\mathbf{v}} \cap \mathbb{C}_{\mathbf{v}'} = \mathbb{R}$;
(5) $\widehat{\mathbf{u}} \not\asymp \widehat{\mathbf{u}}' \Rightarrow \mathbb{S}^1_{\widehat{\mathbf{u}}} \cap \mathbb{S}^1_{\widehat{\mathbf{u}}'} = \{+1, -1\}$.
(6) $\mathsf{q} = q_R + \mathbf{q}_P$, $p = p_R + \mathbf{p}_P$ commute $\Leftrightarrow \mathbf{q}_P \asymp \mathbf{p}_P$.[XV]

SOLUTION

Exercise 3.45. Suppose that $\mathsf{q} \in \mathbb{S}^3$ with $\mathsf{q} := \mathbf{Re}\,\mathsf{q}$. Prove that

(1) $\mathsf{q}^2 = 2q_R \cdot \mathsf{q} - 1$.

(2) $\mathbf{Re}\,\mathsf{q}^2 = 2{q_R}^2 - 1$. Where does this second equation look familiar?

SOLUTION

Exercise 3.46. Show that

$$\frac{2}{5} + \frac{1}{5}\mathbf{i} + \frac{2}{5}\mathbf{j} - \frac{4}{5}\mathbf{k}$$

is an element of $\mathbb{S}^3$. Find its inverse and use Exercise 3.45 to compute its square.

SOLUTION

Since any $\widehat{\mathbf{u}} \in \mathbb{S}^2$ satisfies

$$\widehat{\mathbf{u}}^2 = \mathbf{v}(-\overline{\widehat{\mathbf{u}}}) = -\widehat{\mathbf{u}}\overline{\widehat{\mathbf{u}}} = -\left\|\widehat{\mathbf{u}}\right\|^2 = -1$$

we can see that for fixed $\widehat{\mathbf{u}} \in \mathbb{S}^2$ the set of quaternions

$$\mathbb{C}_{\widehat{\mathbf{u}}} = \{a + b\widehat{\mathbf{u}} : a, b \in \mathbb{R}\}$$

[XIII] Actually there is one complex plane, $\mathbb{C}$, and one circle, $\mathbb{S}^1$, for every pair of anti-polar points in $\mathbb{S}^2$. The set of anti-polar points in $\mathbb{S}^2$ is called the *projective plane* and is denoted as $\mathbb{R}\mathrm{P}^2$. Surprisingly $\mathbb{R}\mathrm{P}^2$ can be made into a surface (i.e. if you lived on it, like on earth, you wouldn't be able to tell that you weren't really on the plane) which cannot be seen as a subset of $\mathbb{R}^3$ without intersecting itself. Check out

en.wikipedia.org/wiki/Boy's_surface

for a description of the so called "Boy's Surface" which shows about as close as we can get to sticking this thing in $\mathbb{R}^3$.

behave exactly the same as the set of complex numbers

$$\{a + b\,i : a, b \in \mathbb{R}\}.$$

In particular we have (**check this!**)

THEOREM 3.107. $(\cos x + \sin x\,\widehat{\mathbf{u}})(\cos y + \sin y\,\widehat{\mathbf{u}}) = \cos(x+y) + \sin(x+y)\,\widehat{\mathbf{u}}$. ■

and inductively

COROLLARY 3.108. $(\cos x + \sin x\,\widehat{\mathbf{u}})^n = \cos nx + \sin nx\,\widehat{\mathbf{u}}$. ■

As such it seems reasonable to define the quaternion

$$e^{\theta,\widehat{\mathbf{u}}} \text{ or } \exp(\theta, \widehat{\mathbf{u}}) = \cos\theta + \sin\theta\,\widehat{\mathbf{u}}.$$

You should be warned that for distant elements of $\widehat{\mathbf{u}}, \widehat{\mathbf{v}} \in \mathbb{S}^2$ we have in general

$$e^{\theta,\widehat{\mathbf{u}}} * e^{\theta,\widehat{\mathbf{v}}} \neq e^{\theta,(\widehat{\mathbf{u}}+\widehat{\mathbf{v}})}.$$

However, all of the usual properties for the exponential carry over to this case so long as we fix which $\widehat{\mathbf{u}} \in \mathbb{S}^2$ we are dealing with, including

- $\left(e^{\theta,\widehat{\mathbf{u}}}\right)^n = e^{(n\theta),\widehat{\mathbf{u}}}$,
- $e^{\theta,\widehat{\mathbf{u}}} * e^{\psi,\widehat{\mathbf{u}}} = e^{(\theta+\psi),\widehat{\mathbf{u}}}$.

The reason why these two equalities holds while the other does not is because both of these equalities occur in a single copy of the circle, $\mathbb{S}^1_{\widehat{\mathbf{u}}}$, while the first involves two different circles, $\mathbb{S}^1_{\widehat{\mathbf{u}}}$ and $\mathbb{S}^1_{\widehat{\mathbf{v}}}$.

All of this allows for rapid calculations of certain powers. For instance, let

$$\omega = \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}. \tag{31}$$

We first observe that $\omega \in \mathbb{S}^3$ since

$$\begin{aligned}\|\omega\|^2 &= \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 \\ &= \frac{2}{4} + \frac{2}{4} + \frac{2}{4} + \frac{2}{4} \\ &= \frac{4}{4} \\ &= 1.\end{aligned}$$

Next, we know that $\cos\frac{\pi}{3} = \frac{1}{2}$ therefore

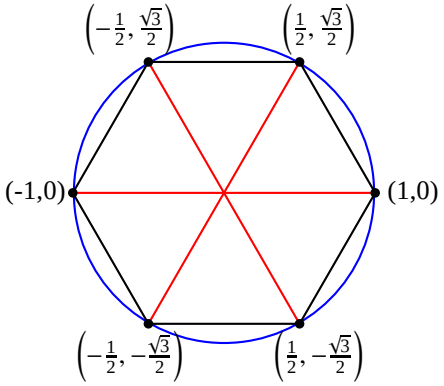
$$\omega = \cos\frac{\pi}{3} + \sin\frac{\pi}{3}\widehat{\mathbf{u}}$$

for some $\widehat{\mathbf{u}} \in \mathbb{S}^2$.[XVI] Drawing out $\{\{6\}\} \subseteq \mathbb{S}^1$:

[XVI] This $\widehat{\mathbf{u}}$ turns out to be

$$\frac{1}{\sqrt{3}}(\mathbf{i} + \mathbf{j} + \mathbf{k}),$$

so that $\omega = \exp(\pi/3 \cdot \frac{1}{\sqrt{3}}(\mathbf{i} + \mathbf{j} + \mathbf{k}))$ but for this calculation that is irrelevant.



The only really important thing to note are the sign changes involved.

$$\begin{aligned}
\omega &= \cos\frac{\pi}{3} + \sin\frac{\pi}{3}\widehat{\mathbf{u}} = \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k} \\
\omega^2 &= \cos\frac{2\pi}{3} + \sin\frac{2\pi}{3}\widehat{\mathbf{u}} = -\cos\frac{\pi}{3} + \sin\frac{\pi}{3}\mathbf{v} = -\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k} \\
\omega^3 &= \cos\frac{3\pi}{3} + \sin\frac{3\pi}{3}\widehat{\mathbf{u}} = \cos\pi + \sin\pi\,\widehat{\mathbf{u}} = -1 \\
\omega^4 &= -\left(\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}\right) = -\frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k} \\
\omega^5 &= \frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k} \\
\omega^6 &= 1.
\end{aligned}$$

Having calculated these powers we can no calculate any power of ω. For instance, since $1,377 = 229 \cdot 6 + 3$ we see that

$$\omega^{1377} = \left(\omega^6\right)^{229} * \omega^3 = 1^{229} \cdot \omega^3 = 1 \cdot \omega^3 = \omega^3 = -1.$$

Actually, the power of ω^n only depends on the class of n mod 6 (in the next section we talk more about this). In other words, if $n = r$ mod 6, then $\omega^n = \omega^r$.

We can also use this to calculate powers of real multiplies of ω as long as we keep track of the power of the real number in question. For example, the quaternion

$$1 + \mathbf{i} + \mathbf{j} + \mathbf{k}$$

is $2 \cdot \omega$. Therefore, if we raise it to the 5^{th} power we get

$$2^5 \cdot \omega^5 = 32 \cdot \left(\frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k}\right) = 16 - 16\mathbf{i} - 16\mathbf{j} - 16\mathbf{k}.$$

EXERCISE 3.47. Calculate

SOLUTION

- ω^{4115} (where ω was defined above, in (31))
- $\left(\frac{1}{4}+\frac{1}{4}\mathbf{i}+\frac{1}{4}\mathbf{j}+\frac{1}{4}\mathbf{k}\right)^{17}$.

Exercise 3.48. Define

$$\alpha \stackrel{\text{def}}{=} \frac{\sqrt{3}}{2}+\frac{3}{11}\mathbf{i}+\frac{1}{11}\mathbf{j}-\frac{9}{22}\mathbf{k}.$$

First show that $\alpha \in \mathbb{S}^3$. Then make a unit circle which helps you calculate all of the powers of α. In particular, calculate α^{117}.

SOLUTION

Example 3.109. A more complicated quaternion to deal with is

$$\mathfrak{q}_5 := \frac{\phi}{2}+\frac{1}{2}\mathbf{i}+\frac{\phi^{-1}}{2}\mathbf{j}$$

where ϕ is the golden ratio. As part of Exercise 3.35 you show that $\| \mathfrak{q}_5 \| = 1$, i.e. $\mathfrak{q}_5 \in \mathbb{S}^3$. Moreover, since

$$\frac{\phi}{2} = \cos\frac{\pi}{5}$$

we see that

$$\mathfrak{q}_5 = \cos\frac{\pi}{5}+\sin\frac{\pi}{5}\widehat{\mathbf{u}}$$

where, again, $\widehat{\mathbf{u}}$ doesn't really matter.[XVII] Let $a := \sin\frac{\pi}{5}$. Then using basic double angle formulas and identities involving the golden ration we have

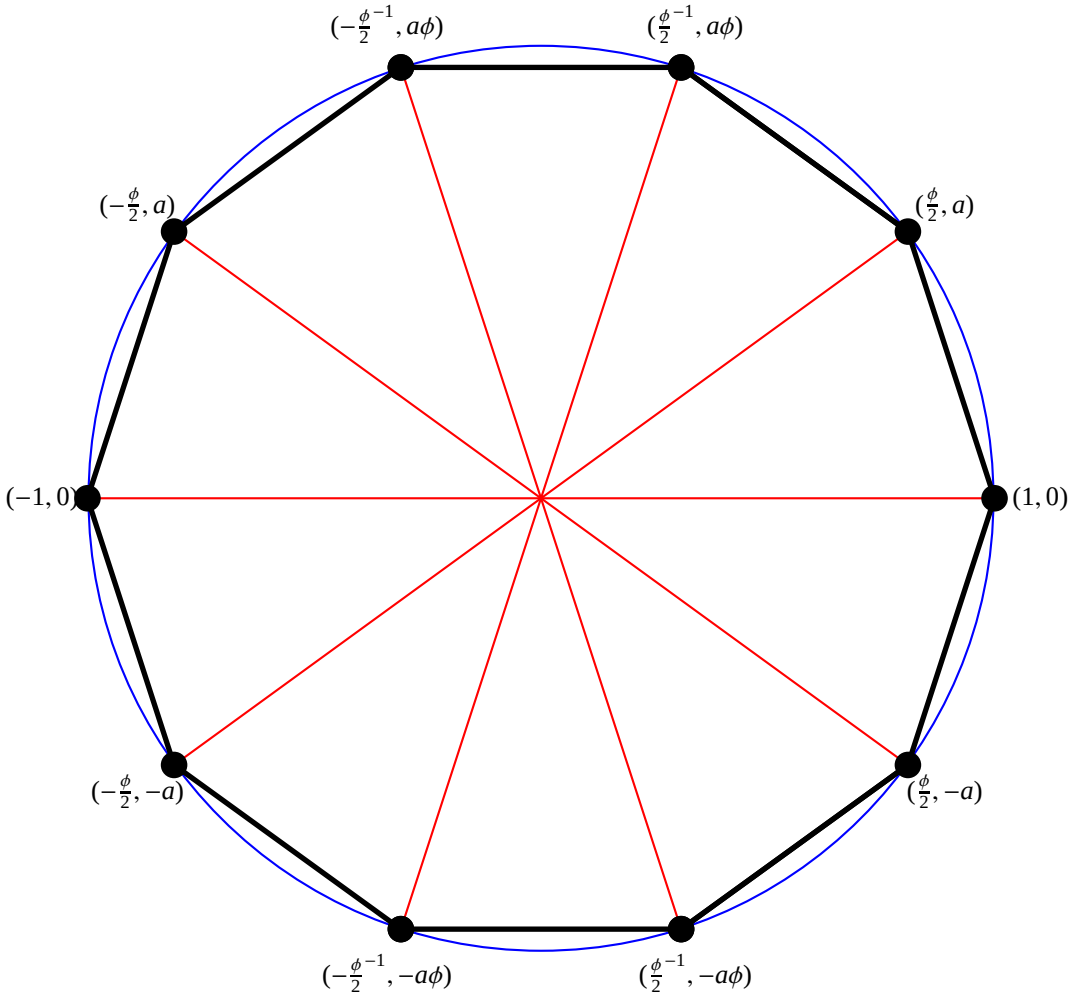
$$2\cos\frac{2\pi}{5} = \phi^2 - 2 = \phi + 1 - 2 = \phi - 1 = \phi^{-1} \Rightarrow \cos\frac{2\pi}{5} = \frac{\phi^{-1}}{2}.$$

At the same time we also have

$$\sin\frac{2\pi}{5} = 2\sin\frac{\pi}{5}\cos\frac{\pi}{5} = \phi a.$$

[XVII] If you're dying to know, as always, $\widehat{\mathbf{u}}$ is just the normalization of the pure part, here

$$\widehat{\mathbf{u}} = \frac{1}{\sqrt{2+\phi}}(\phi\mathbf{i}+\mathbf{j}).$$


$(-\frac{\phi}{2}^{-1}, a\phi)$
$(\frac{\phi}{2}^{-1}, a\phi)$
$(-\frac{\phi}{2}, a)$
$(\frac{\phi}{2}, a)$
$(-1, 0)$
$(1, 0)$
$(-\frac{\phi}{2}, -a)$
$(\frac{\phi}{2}, -a)$
$(-\frac{\phi}{2}^{-1}, -a\phi)$
$(\frac{\phi}{2}^{-1}, -a\phi)$

Using these facts and taking advantage of symmetry allows us to fill in the coordinates of the 10 points of {{10}}. Now we can quickly figure out all powers of q_5:

$$q_5^0 = 1 \qquad q_5 = \frac{\phi}{2} + \frac{1}{2}\mathbf{i} + \frac{\phi^{-1}}{2}\mathbf{j}$$
$$q_5^2 = \frac{\phi^{-1}}{2} + \frac{\phi}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} \qquad q_5^3 = -\frac{\phi^{-1}}{2} + \frac{\phi}{2}\mathbf{i} + \frac{1}{2}\mathbf{j}$$
$$q_5^4 = -\frac{\phi}{2} + \frac{1}{2}\mathbf{i} + \frac{\phi^{-1}}{2}\mathbf{j} \qquad q_5^5 = -1$$
$$q_5^6 = -\frac{\phi}{2} - \frac{1}{2}\mathbf{i} - \frac{\phi^{-1}}{2}\mathbf{j} = -q_5 \qquad q_5^7 = -\frac{\phi^{-1}}{2} - \frac{\phi}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} = -q_5^2$$
$$q_5^8 = \frac{\phi^{-1}}{2} - \frac{\phi}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} = -q_5^3 \qquad q_5^9 = \frac{\phi}{2} - \frac{1}{2}\mathbf{i} - \frac{\phi^{-1}}{2}\mathbf{j} = -q_5^4$$
$$q_5^{10} = 1.$$

and then things begin to cyclically repeat (as before). For instance

$$q_5^{31456} = q_5^6 = -\frac{\phi}{2} - \frac{1}{2}\mathbf{i} - \frac{\phi^{-1}}{2}\mathbf{j} = -q_5.$$

We will see the significance of q_5 in the appendix. See 790.

Remark 3.110. We can think of this unit circle as living in the $\mathbb{C}_{\mathbf{v}}$ plane ... i.e the x-axis is labeled by 1 and the y-axis would be labeled (instead of by i as in the usual complex plane) by $\mathbf{v}$. Geometrically, this is a the intersection of a two dimensional plane with the 4 dimensional space $\mathbb{R}^4$ (or $\mathbb{H}$) and the circle we see is the intersection of this same two dimensional plane with $\mathbb{S}^3$. Such pictures are nice but you could drive yourself crazy thinking of three dimensional spheres in four dimensions and various intersections of it with hyper planes. Many a student has leapt of this cliff never to return. For our purposes the quaternions really shouldn't be thought of geometrically since this provides little to no insight into our use for them! However, if you feel confident enough with the basics, and you are not easily distracted please feel free to google some wonderful ways to kinda-sorta visualize $\mathbb{S}^3$.

Exercise 3.49. Find 3 quaternions, $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, each with more than one term (so $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ don't count!) so that

$$\mathbf{u}^2 = \mathbf{v}^2 = \mathbf{w}^2 = -1, \quad \mathbf{u} * \mathbf{v} = \mathbf{w}.$$

Show that no matter which triple you have which satisfy the above one always has that

$$\mathbf{u} * \mathbf{v} = -\mathbf{v} * \mathbf{u}.$$

SOLUTION

Hint! As the notation suggests, each of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ must be pure. Show that if the product of two pure quaternions is pure then they anti-commute.

The following exercise (partially) motivates the non-commutative nature of quaternions.

Exercise 3.50. The following statement (and hence its proof) is false. Find the error in the proof.

False Theorem 3.111. *The quaternion* $\mathbf{i}+\mathbf{j}$ *is not invertible (i.e. has no inverse).*

"**Proof**". Suppose we have a quaternion q so that

$$\mathsf{q} * (\mathbf{i}+\mathbf{j}) = 1. \tag{32}$$

Then we would have that

$$\begin{aligned}
(\mathbf{i}-\mathbf{j}) &= 1 * (\mathbf{i}-\mathbf{j}) \\
&= (\mathsf{q} * (\mathbf{i}+\mathbf{j})) * (\mathbf{i}-\mathbf{j}) && \text{(Multiply bs of (32) by } (\mathbf{i}-\mathbf{j})) \\
&= \mathsf{q} * ((\mathbf{i}+\mathbf{j}) * (\mathbf{i}-\mathbf{j})) && \text{(Associativity)} \\
&= \mathsf{q} * (\mathbf{i}^2 - \mathbf{j}^2) && \text{(Diff of 2 squares)} \\
&= \mathsf{q} * ((-1) - (-1)) && (\mathbf{i}^2 = \mathbf{j}^2 = -1) \\
&= \mathsf{q} * 0 && ((-1)-(-1) = -1+1 = 0) \\
&= 0.
\end{aligned}$$

Therefore $\mathbf{i}-\mathbf{j}=0$, which is false and so no such q can exist. ■

SOLUTION

Remark 3.112. In the above exercise, if $\mathbf{i}$ and $\mathbf{j}$ commuted (which they don't) the statement would be true, and we would lose the precious ability to divide. In practice, this is a very good trade. In fact, there is an analogue of the quaternions defined in 8 dimensions called the octonions. There we even lose the associative property of triples.[XVIII].

3.3.2 Picturing $\mathbb{S}^3$*

I suggest you skip this section the first time you real through this, but I know you probably won't ☹. Students always try to picture the quaternions and $\mathbb{S}^3$ despite my warnings. So, like a parent who gives their kid their first beer so they avoid them running off to some crazy part to try it, I will quickly explain a couple of ways of explaining how to picture $\mathbb{S}^3$. Now $\mathbb{S}^3$ is a curved surface in $\mathbb{R}^4$ (the same way that $\mathbb{S}^2$ is a curved surface in $\mathbb{R}^3$) so it is impossible to see directly. There are however a couple of standard tricks to attempt to see this.

The first is to make a "movie" with frames (for t varying in the real numbers)

$$\mathbb{R}^3_t \overset{\text{def}}{=} t^{\text{th}} \text{ frame in movie} = \{t + a\mathbf{i} + b\mathbf{j} + c\mathbf{k} : a,b,c \in \mathbb{R}\}$$
$$\mathbb{S}^2_\theta \overset{\text{def}}{=} \mathbb{R}^3_{\cos\theta} \cap \mathbb{S}^3 = \{\cos\theta + \sin\theta\,\widehat{\mathbf{u}} : \widehat{\mathbf{u}} \in \mathbb{S}^2\}.$$

Note that

$$\mathbb{S}^2_0 = \{1\} \quad \text{and} \quad \mathbb{S}^2_\pi = \{+1\}.$$

Note that $\mathbb{S}^2_\theta$ is (thinking of the coordinates of $\mathbb{H}$ as (t,x,y,z))

$$\cos\theta^2 + x_1^2 + x_2^2 + x_3^2 = 1 \Rightarrow x_1^2 + x_2^2 + x_2^2 = 1 - \cos\theta^2 = \sin^2\theta$$

a sphere of radius $\sin\theta$. In analogy with lines of latitude, these guys foliate $\mathbb{S}^3$ into a sequence of spheres which get smaller the same way starting from $\theta = \pi/2$. On the other hand, the circles, $\mathbb{S}^1_{\widehat{\mathbf{u}}}$, like lines of longitude on the Earth, do **not** change size. Moreover we have that

$$\mathbb{S}^2_\theta \cap \mathbb{S}^1_{\widehat{\mathbf{u}}} = \{\cos\theta + \sin\theta\,\widehat{\mathbf{u}}, \cos\theta - \sin\theta\,\widehat{\mathbf{u}}\}$$

which are two distinct points for $0 < \theta < \pi$ and and 1 for $\theta = 0$ or $\theta = \pi$. Again, this is exactly what happens with lines of latitude and longitude on Earth. In terms of group theory (we haven't introduced these terms yet) the lines of latitude, the $\mathbb{S}^2_\theta$'s, will make up the conjugacy classes of $\mathbb{S}^3$ and these lines of longitude, the $\mathbb{S}^1_{\widehat{\mathbf{u}}}$, make up the centralizers of $\mathbb{S}^3$.

One way of thinking of $\mathbb{S}^2$ is as the union of the northern and southern hemispheres glued together along to the equation. The advantage of this is that we can flatten out these two hemispheres in the plane. In this way, lines of latitude become concentric circles on the discs, while lines of longitude become a straight line on each each disc which meet the boundary of each disc into a pair of points which end up getting glued together. Motivated from the picture that allows us to go from 3 to 2 dimensions, we can move from 3 dimensions to 2 dimensions to see $\mathbb{S}^3$ as two solid balls glued together along the boundary. Unlike the case of two dimensions, we cannot see the resulting space since when glueing things together, we end up in four dimensions. Indeed, the two disks are

$$B_+ := \{(t,x,y,z) : t^2 + x^2 + y^2 + z^2 = 1, t \geq 0\} \quad \text{and} \quad B_- := \{(t,x,y,z) : t^2 + x^2 + y^2 + z^2 = 1, t \leq 0\}.$$

As promised, $B_+ \cap B_- = \mathbb{S}^2$, i.e.

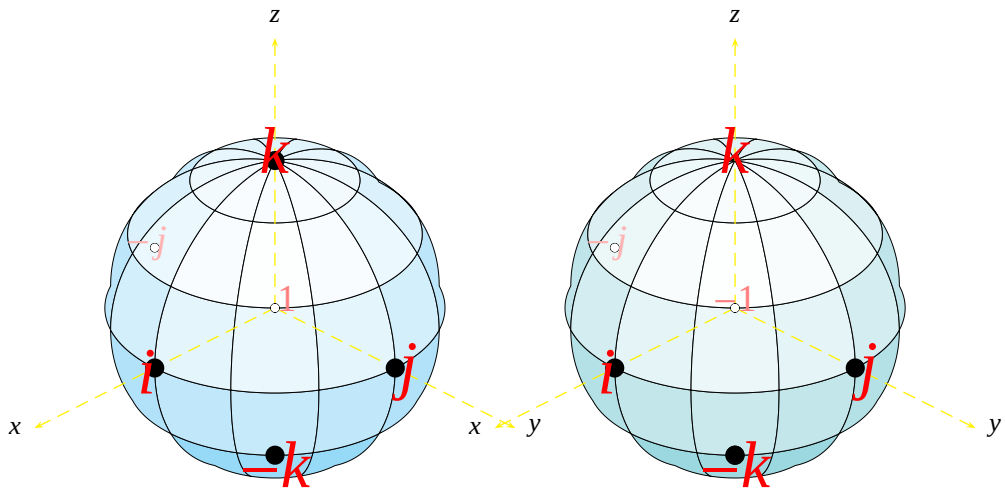
$$\{(t,x,y,z) : x^2 + y^2 + z^2 = 1, t \geq 0, t \leq 0\} = \{(t,x,y,z) : x^2 + y^2 + z^2 = 1, t = 0\} = \{(0,x,y,z) : x^2 + y^2 + z^2 = 1\} = \mathbb{S}^2.$$

Moreover, each half can be identified with a solid ball, $\{(x,y,z) : x^2 + y^2 + z^2 \leq 1\}$ in $\mathbb{R}^3$ via

$$(t,x,y,z) \mapsto (x,y,z)$$

which have inverses

$$(x,y,z) \mapsto (\sqrt{1 - (x^2 + y^2 + z^2)}, x, y, z) \in B_+ \quad \text{and} \quad (x,y,z) \mapsto (-\sqrt{1 - (x^2 + y^2 + z^2)}, x, y, z) \in B_-.$$

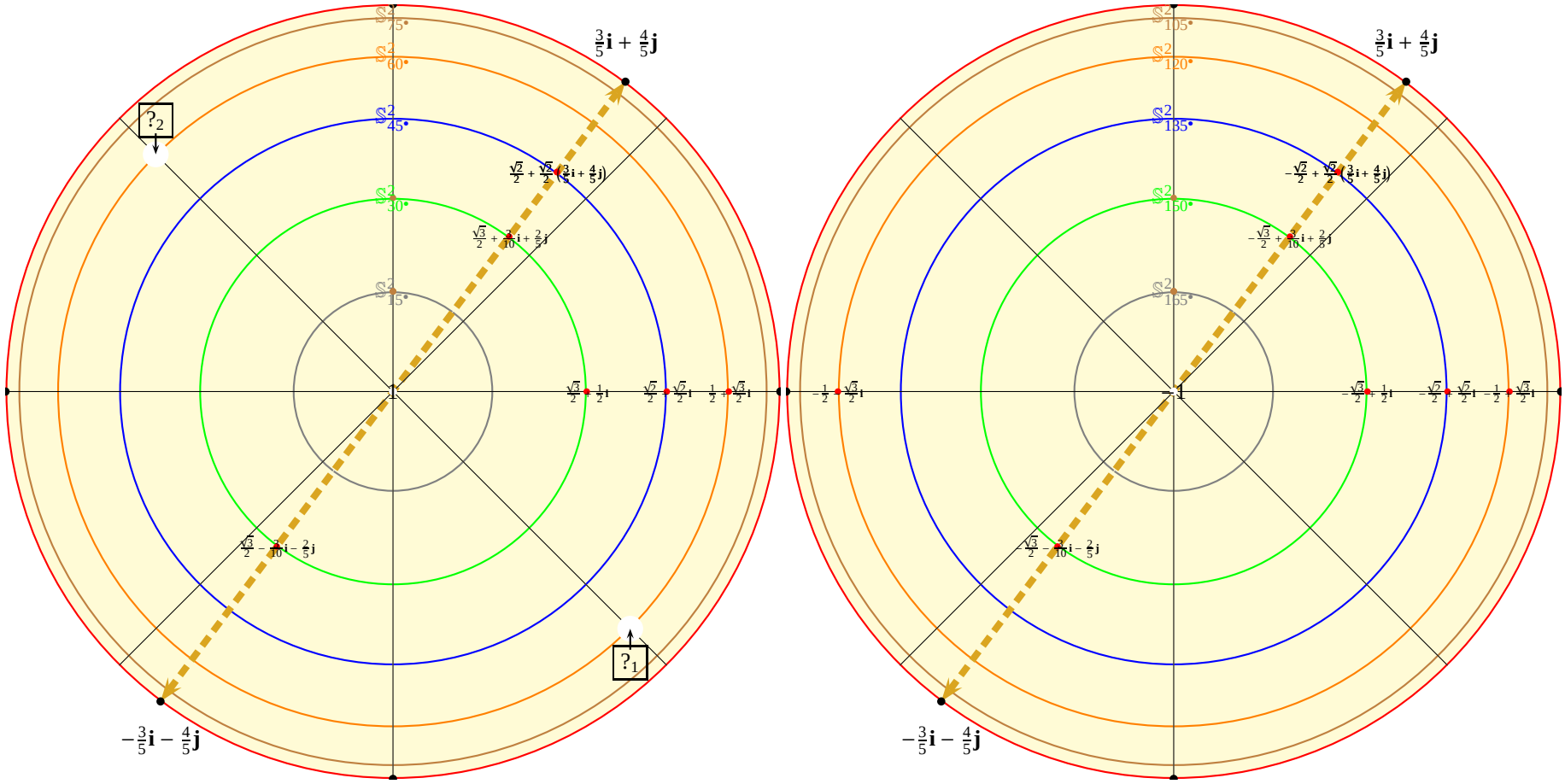


$0 \leq \theta \leq \pi/2$ $\pi/2 \leq \theta \leq \pi$

In this way latitudes, $\mathbb{S}^2_\theta$, correspond to shells inside of one of the two spheres, if $0 < \theta < \pi/2$ it's a shell inside of the left hand sphere and if $\pi/2 < \theta < \pi$ it's a shell inside of the right hand sphere. The set of pure unit quaternions, $\mathbb{S}^2 = \mathbb{S}^2_{\pi/2}$ appear in both spheres as the outside shell. The two balls are to be glued together along this outer sphere. This is analogous to a 2 sphere, for instance the surface of the Earth, can be obtained by glueing together the northern and souther hemispheres along the equator, a circle. Therefore the sphere on the left hand side can be thought of as "the northern hemisphere" and ball on the right hand side as the "southern hemisphere", and the set of pure unit quaternions as the equator. The lines of latitude, based on the way we've drawn things, appear to be straight lines: they are the straight lines.[XIX] Indeed, the longitude $\mathbb{S}^1_{\widehat{\mathbf{u}}}$ goes though the origin of both spheres, and connect the points $\widehat{\mathbf{u}}$ and $-\widehat{\mathbf{u}}$ on the pure shells.

[XIX] Indeed, if we stand at the north pole and look at the lines of latitude on Earth, they too will appear to be straight lines.

Here we draw a cross section of these two spheres (looking at quaternions in $\mathbb{S}^3$ with $\mathbf{k}$ component 0). In the picture the concentric circles are the latitude lines, i.e. the $\mathbb{S}^2_\theta$'s and the straight line, in gold, is the longitude $\mathbb{S}^1_{\frac{3}{5}\mathbf{i}+\frac{4}{5}\mathbf{j}}$



What are the points with the question marks? Well they're located on $\mathbb{S}^2_{60^\circ}$ and $\mathbb{S}^1_{\frac{\sqrt{2}}{2}\mathbf{i}-\frac{\sqrt{2}}{2}\mathbf{j}} = \mathbb{S}^1_{-\frac{\sqrt{2}}{2}\mathbf{i}+\frac{\sqrt{2}}{2}\mathbf{j}}$ hence it must be

$$?_1 = \cos 60^\circ + \sin 60^\circ \left(\frac{\sqrt{2}}{2}\mathbf{i} - \frac{\sqrt{2}}{2}\mathbf{j}\right) = \frac{1}{2} + \frac{\sqrt{6}}{4}\mathbf{i} - \frac{\sqrt{6}}{4}\mathbf{j} \qquad ?_2 = \frac{1}{2} - \frac{\sqrt{6}}{4}\mathbf{i} + \frac{\sqrt{6}}{4}\mathbf{j}.$$

3.3.3 Collection of Facts

This table collects all the basic facts about quaternions:

(1) First we present three formulas which can be used to define quaternionic multiplication (depending on whether we think of $\mathbb{H}$ as $\mathbb{R}^4$, $\mathbb{C}^2$, or $\mathbb{R} \times \mathbb{R}^3$).

($\mathbb{R}^4$ **Version**)

$$\begin{aligned}(a + b\,\mathbf{i} + c\,\mathbf{j} + d\,\mathbf{k}) * (x + y\,\mathbf{i} + z\,\mathbf{j} + w\,\mathbf{k}) \\ = (ax - by - cz - dw) + (bx + ay + dz - cw)\,\mathbf{i} \\ + (cx + az + bw - dy)\,\mathbf{j} + (aw + dx + cy - bz)\,\mathbf{k}.\end{aligned}$$

($\mathbb{C} \times \mathbb{C}$ **Version**)

$$(z_1, z_2) * (w_1, w_2) = (z_1 w_1 - z_2 \overline{w}_2, z_1 w_2 + z_2 \overline{w}_1)$$

($\mathbb{R} \times \mathbb{R}$ **Version**)

$$(a, \mathbf{v}) * (b, \mathbf{w}) = (ab - \mathbf{v} \bullet \mathbf{w}, a\mathbf{w} + b\mathbf{w} + \mathbf{v} \times \mathbf{w})$$ [xx]

(2) Here are some formulas related to the conjugate.

- $\overline{q + p} = \overline{q} + \overline{p}$
- For all $a \in \mathbb{R}$, $\overline{a} = a$
- For all $\mathbf{v} \in \mathbb{R}^3$ $\overline{\mathbf{v}} = -\mathbf{v}$
- $\overline{aq} = a\overline{q}$
- When $q \in \mathbb{S}^3$ (identified with quaternions whose length is 1) we have
$$\overline{q} = q^{-1}$$
- $\overline{q} * q = q * \overline{q} = \| q \|^2$
- $\overline{q * p} = \overline{p} * \overline{q}$ (notice the reverse!)

(3) A few more general facts/ formulas

- When $q \in \mathbb{S}^2$ (identified with pure quaternions whose length is 1) we have
$$q^2 = -1$$
- $(q * p)^{-1} = p^{-1} * q^{-1}$ (Again, notice the reverse!)
- $\| q * p \| = \| q \| \cdot \| p \|$.

[xx] We also use this formula after replacing $(a, \mathbf{v})$ with $a + \mathbf{v}$.

3.3.4 More Quaternion Odds and Ends

Example 3.113. In the quaternions there seems there is always an easier way to do a given calculation than the way you are doing it.

(1) Let's calculate

$$\omega * \mathbf{i} * \omega^{-1}$$

where ω is defined as in Equation (31) on page 128.

(a) (**Long Way**) For starters, $\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k} \in \mathbb{S}^3$ so its inverse is just its conjugate: $\frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k}$. Moreover we have that

$$\begin{aligned}\left(\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}\right) * \mathbf{i} &= \frac{\mathbf{i} + (-1) + (-\mathbf{k}) + \mathbf{j}}{2} \\ &= -\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k},\end{aligned}$$

so

$$\begin{aligned}\left(\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}\right) * \mathbf{i} * \left(\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}\right)^{-1} &= \left(-\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k}\right) * \left(+\frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k}\right) \\ &= \frac{(-1 + \mathbf{i} + \mathbf{j} - \mathbf{k}) * (1 - \mathbf{i} - \mathbf{j} - \mathbf{k})}{4}\end{aligned}$$

Using the properties of the cross product in $\mathbb{R}^3$ saves us some time. So if we let $\mathbf{v} := \mathbf{i} + \mathbf{j} - \mathbf{k}$ and $\mathbf{w} := -\mathbf{i} - \mathbf{j} - \mathbf{k}$ we have that $\mathbf{v} \bullet \mathbf{w} = -1 - 1 + 1 = -1$ and

$$\begin{aligned}\mathbf{v} \times \mathbf{w} &= (\mathbf{i} + \mathbf{j} - \mathbf{k}) \times (-\mathbf{i} - \mathbf{j} - \mathbf{k}) \\ &= -((\mathbf{i} + \mathbf{j}) - \mathbf{k}) \times ((\mathbf{i} + \mathbf{j}) + \mathbf{k}) \\ &= -(\mathbf{i} + \mathbf{j}) \times \mathbf{k} + \mathbf{k} \times (\mathbf{i} + \mathbf{j}) \\ &= -(\mathbf{i} \times \mathbf{k}) - (\mathbf{j} \times \mathbf{k}) + (\mathbf{k} \times \mathbf{i}) + (\mathbf{k} \times \mathbf{j}) \\ &= (\mathbf{k} \times \mathbf{i}) + (\mathbf{k} \times \mathbf{j}) + (\mathbf{k} \times \mathbf{i}) + (\mathbf{k} \times \mathbf{j}) \\ &= 2(\mathbf{k} \times \mathbf{i}) + 2(\mathbf{k} \times \mathbf{j}) \\ &= 2\mathbf{j} - 2\mathbf{i}.\end{aligned}$$

So using the formula above we get that

$$\begin{aligned}\frac{(-1 + \mathbf{i} + \mathbf{j} - \mathbf{k}) * (1 - \mathbf{i} - \mathbf{j} - \mathbf{k})}{4} &= \frac{-1 - \mathbf{v} \bullet \mathbf{w} + (-1)\mathbf{w} + (1)\mathbf{v} + \mathbf{v} \times \mathbf{w}}{4} \\ &= \frac{0 + (\mathbf{i} + \mathbf{j} + \mathbf{k}) + (\mathbf{i} + \mathbf{j} - \mathbf{k}) + 2\mathbf{j} - 2\mathbf{i}}{4} \\ &= \frac{4}{4}\mathbf{j} \\ &= \mathbf{j}.\end{aligned}$$

(b) (**Short Way**) On the other hand, if we were lucky enough to notice that

$$\frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{k}=\left(\frac{\sqrt{2}}{2}(1+\mathbf{k})\right)*\left(\frac{\sqrt{2}}{2}(1+\mathbf{i})\right)=\frac{1}{2}(1+\mathbf{k})*(1+\mathbf{i})$$

this problem becomes a lot easier. Indeed

$$\begin{aligned}
\omega*\mathbf{i}*\omega^{-1} &= \left(\frac{\sqrt{2}}{2}(1+\mathbf{k})*\frac{\sqrt{2}}{2}(1+\mathbf{i})\right)*\mathbf{i}*\left(\frac{\sqrt{2}}{2}(1+\mathbf{k})*\frac{\sqrt{2}}{2}(1+\mathbf{i})\right)^{-1} && (33)\\
&= \left(\frac{\sqrt{2}}{2}(1+\mathbf{k})\right)*\left(\frac{\sqrt{2}}{2}(1+\mathbf{i})\right)*\mathbf{i}*\left(\frac{\sqrt{2}}{2}(1+\mathbf{i})\right)^{-1}*\left(\frac{\sqrt{2}}{2}(1+\mathbf{k})\right)^{-1} && (34)\\
&= \left(\frac{\sqrt{2}}{2}(1+\mathbf{k})\right)*\mathbf{i}*\left(\frac{\sqrt{2}}{2}(1+\mathbf{k})\right)^{-1} && (35)\\
&= \left(\frac{\sqrt{2}}{2}(1+\mathbf{k})\right)^{2}*\mathbf{i} && (36)\\
&= \mathbf{k}*\mathbf{i} && (37)\\
&= \mathbf{j}. && (38)
\end{aligned}$$

In this calculation we made liberal use of the fact that if $\mathbf{Pu}\,\mathsf{q}\asymp\mathbf{v}$, then $\mathsf{q}*\mathbf{v}*\mathsf{q}^{-1}=\mathbf{v}$ to go from (34) to (35). Then I used the fact that if $\mathbf{Pu}\,\mathsf{q}\perp\mathbf{v}$, then $\mathsf{q}*\mathbf{v}*\mathsf{q}^{-1}=\mathsf{q}^{2}*\mathbf{v}$(35) to (36). To go from line (33) to (34) I used a general principle known as 'Socks and Shoes,' which will be proven in general in Theorem 4.37. It says that

$$(\mathsf{q}*\mathsf{p})^{-1}=\mathsf{p}^{-1}*\mathsf{q}^{-1}.$$

(2) We could also see this same element

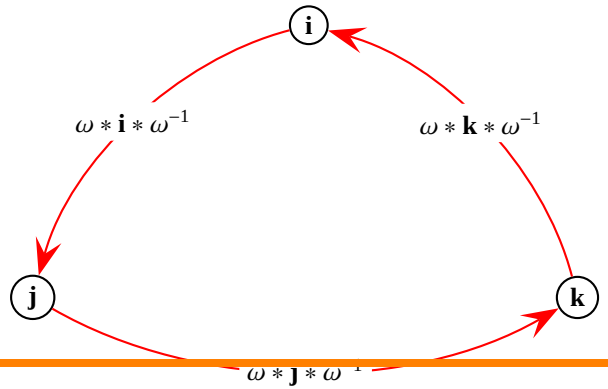
$$\omega=\frac{1}{2}\left(1+\mathbf{i}\right)*\left(1+\mathbf{j}\right)$$

and this would allow us to easily calculate

$$\omega*\mathbf{j}*\omega^{-1}=\mathbf{k}.$$

This calculation is simple again because the $1+\mathbf{j}$ term commutes with the $\mathbf{j}$ while the $1+\mathbf{i}$ anti-commutes with term stuck in the middle.

(3) Here's a sweet diagram to help you remember what the effect of $\omega*\mathsf{q}*\omega^{-1}$ is.



PROPOSITION 3.114. *Any quaternion* q *can be written as the product of two pure quaternions. Moreover if* $\mathsf{q} \in \mathbb{S}^3$*, then both pure quaternions can also be taken to have length 1 (hence are elements of* $\mathbb{S}^2$*).*

PROOF. Let q be a quaternion. The result is easy if q is real, so we assume that q is not real. That is $\mathbf{Pu}(\mathsf{q})$ is non-zero. Select a unit vector, i.e. an element of $\mathbb{S}^2$, $\widehat{\mathbf{u}}$, which is perpendicular to $\mathbf{Pu}(\mathsf{q})$. Since

$$\mathsf{q} * \widehat{\mathbf{u}} = (-\mathbf{Pu}(\mathsf{q}) \bullet \widehat{\mathbf{u}},\ \mathbf{Re}(\mathsf{q})\widehat{\mathbf{u}} + \mathbf{Pu}(\mathsf{q}) \times \widehat{\mathbf{u}}) = (0, \mathbf{Re}(\mathsf{q})\widehat{\mathbf{u}} + \mathbf{Pu}(\mathsf{q}) \times \widehat{\mathbf{u}}),$$

where we've used the fact that $\mathbf{Pu}(\mathsf{q}) \perp \widehat{\mathbf{u}}$, the product $\mathsf{q} * \widehat{\mathbf{u}}$ is a pure quaternion, say $\mathbf{v}$. Multiplying both sides of our new equation

$$\mathsf{q} * \widehat{\mathbf{u}} = \mathbf{v}$$

by $-\widehat{\mathbf{u}}$, which happens to be $\widehat{\mathbf{u}}^{-1}$ since $\widehat{\mathbf{u}} \in \mathbb{S}^2$, we see that

$$\mathsf{q} = \mathbf{v} * \widehat{\mathbf{u}}' \tag{39}$$

where $\widehat{\mathbf{u}}' := -\widehat{\mathbf{u}}$ is pure. We have now proven that any $\mathsf{q} \in \mathbb{H}$ can be written as the product of two pure quaternions.

Now if $\mathsf{q} \in \mathbb{S}^3$ and since $\widehat{\mathbf{u}} \in \mathbb{S}^2$ and hence also in $\mathbb{S}^3$ by assumption on $\widehat{\mathbf{u}}$, we know that the equation

$$\mathsf{q} * \widehat{\mathbf{u}} = \mathbf{v}$$

implies that $\mathbf{v} \in \mathbb{S}^3$ by closure of multiplication in $\mathbb{S}^3$ and in this special case we have proven that we can take both factors in (39) to be in $\mathbb{S}^2$. ■

EXAMPLE 3.115. Let's follow the proof to factor ω from (31) as a product of two elements of $\mathbb{S}^2$. The pure part, $\mathbf{Pu}(\omega)$ is

$$\frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}.$$

A unit vector perpendicular to this vector could be taken (among other things) to be

$$\widehat{\mathbf{u}} := \frac{\sqrt{2}}{2}(\mathbf{i} - \mathbf{j})$$

which we denote as $\widehat{\mathbf{u}}$. Now

$$\omega * \widehat{\mathbf{u}} = \mathbf{Re}(\omega) \cdot \widehat{\mathbf{u}} + \mathbf{Pu}(\omega) \times \widehat{\mathbf{u}} = \frac{\sqrt{2}}{2}(\mathbf{i} - \mathbf{k}).$$

And accordingly we have written ω as the product of two pure quaternions:

$$\begin{aligned}
\omega &= \left(\frac{\sqrt{2}}{2}(\mathbf{i} - \mathbf{k})\right) * \widehat{\mathbf{u}}^{-1} \\
&= \left(\frac{\sqrt{2}}{2}(\mathbf{i} - \mathbf{k})\right) * (-\widehat{\mathbf{u}}^{-1}) \\
&= \left(\frac{\sqrt{2}}{2}(\mathbf{i} - \mathbf{k})\right) * \left(\frac{\sqrt{2}}{2}(\mathbf{j} - \mathbf{i})\right).
\end{aligned}$$

NOTATION 3.116. For every $\widehat{\mathbf{u}} \in \mathbb{S}^2$ define

$$S_{\widehat{\mathbf{u}}} = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\widehat{\mathbf{u}}.$$

Note that

$$\| S_{\widehat{\mathbf{u}}} \|^2 = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = 1 \Rightarrow S_{\widehat{\mathbf{u}}} \in \mathbb{S}^3.$$

Moreover

$$e^{\frac{\pi}{4}, \widehat{\mathbf{u}}} = S_{\widehat{\mathbf{u}}}.$$

In particular, $S_{\widehat{\mathbf{u}}}^2 = e^{\frac{\pi}{2}, \widehat{\mathbf{u}}} = \widehat{\mathbf{u}}$, $S_{\widehat{\mathbf{u}}}^4 = -1$, and $S_{\widehat{\mathbf{u}}}^8 = 1$.

EXERCISE 3.51. Prove that if $\widehat{\mathbf{u}}$ and $\widehat{\mathbf{v}}$ are perpendicular elements of $\mathbb{S}^2$, then

$$S_{\widehat{\mathbf{u}}} * e^{\theta, \widehat{\mathbf{v}}} * S_{\widehat{\mathbf{u}}}^{-1} = e^{\theta, \widehat{\mathbf{u}} \times \widehat{\mathbf{v}}}.$$

SOLUTION

Since $e^{\frac{\pi}{4}, \widehat{\mathbf{u}}} = S_{\widehat{\mathbf{u}}}$, a particular special case of this exercise states

$$S_{\widehat{\mathbf{u}}} * S_{\widehat{\mathbf{u}}'} * S_{\widehat{\mathbf{u}}}^{-1} = S_{\widehat{\mathbf{u}} \times \widehat{\mathbf{u}}'}.$$

3.4 Modular Arithmatic

To everything – turn, turn, turn

We have spent a while studying the quaternions. This set continued our chain

$$\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C} \subseteq \mathbb{H}.$$

(we didn't study the rationals or the reals). Now we take things in the complete opposite direction and study something that is, in a sense, *smaller* than the integers. Recall that for every positive integer n we defined a relation $\equiv_n$ by $a \equiv_n b$ iff $a - b$ is a multiple of n (i.e. if n divides $a - b$). We showed that this was an equivalence relation. Like every equivalence relation we could cut the integers up into class. The class of an integer m was denoted $[\, m \,]_n$. The set of all classes was denoted $\mathbb{Z}_n$. No set could possibly feel less like the quaternions! Here is the most obvious difference

- The quaternions $\mathbb{H}$ are uncountably infinite.[I] In fact, as a set, they are (in bijection with) $\mathbb{R}^4$.
- We will see that $\mathbb{Z}_n$ has only n elements. In particular $\mathbb{Z}_n$ is always finite!

Despite this we will ask very similar questions including

- How do we add elements of $\mathbb{Z}_n$?
- How do we multiply elements of $\mathbb{Z}_n$?
- Are these operations associative?
- Are the operations commutative?
- Which elements have a multiplicative inverse? For such elements is there a simple formula for it?

3.4.1 Definitions and Basics

Here are some intuitive examples:

Remember $a \equiv_n b \Leftrightarrow [a]_n = [b]_n$.

Example 3.117. Modular arithmetic is something that we have done many times without realizing it. For instance

(1) Today is Tuesday. If I wanted to figure out what day of the week it would be in 2146 days I would note that

$$2148 = 300 \cdot 7 + 46 = 300 \cdot 7 + 300 \cdot 7 + 6 \cdot 7 + 4 = 306 \cdot 7 + 6 = 307 \cdot 7 - 1.$$

Hence 2146 days from today would be a Monday.

(2) It is 3AM right now. To figure out what time it would be 5282 hours from now I first note that

$$5282 = 220 \cdot 24 + 2$$

so the clock would show 2 hours more and it would be 5AM.

[I] Meaning they have the same cardinality as $\mathbb{R}$ does.

(3) I have a number x which ends in 7 (e.g 7, 207, or 801,029,423,905,902,349,082,417) and I add to it $y = 1,024,181,293,142,895$ (or any other integer ending in 5). Then since the 1's place of this second number is 5 I know the 1's place of the sum $x + y$ is 2 (this is because $7 + 5$ has one's place 2.

(4) I'm at a a 15 degree angle (w.r.t. the x axis) and I rotate by 2013902453 degrees. Then since $2013902453 = 5594173 \cdot 360 + 173$ I know that I would be at $173 + 15 = 188$ degrees.

(5) (Baseball Example!!!) The Red Sox's starting rotation for the year 2011 was Jon Lester, Clay Buchholz, John Lackey, Josh Beckett, and Daisuke Matsuzaka as the fifth starter. (I know, Lackey before Beckett is a joke!). Assuming nobody misses a start (another joke) the starter of the final game of the season (a baseball season has 162 games) would be Clay Buchholz's.

(6) It's summer right now. Seventeen seasons from now it will be Fall. This is because

$$17 = 16 + 1 = 4 \cdot 4 + 1.$$

In Example 2.72.2 back on page 50 we showed that the relation

$$a \equiv_n b \Leftrightarrow n \mid (b - a)$$

was an equivalence relation. Later, on page 56 we agreed to denote the class of a as $[\, a \,]_n$ and the set of all such classes as $\mathbb{Z}_n$. Notice that

$$\begin{aligned} [\, a \,]_n &= \{b : a \equiv_n b\} \\ &= \{b : n \mid (b - a)\} \\ &= \{b : qn = b - a \text{ for some } q \in \mathbb{Z}\} \\ &= \{b : b = a + qn \text{ for some } q \in \mathbb{Z}\}. \end{aligned}$$

This set was previously denoted as $n\mathbb{Z} + \{a\}$ (or as $n\mathbb{Z} + a$) in Equation (2) back on page 75. Therefore we have

THEOREM 3.118. *For any integers a and n we have*

$$[\, a \,]_n = n\mathbb{Z} + a.$$

■

CONVENTION 3.119. Since it's less writing, we will still typically denote the class of a mod n as $[\, a \,]_n$ and not as $n\mathbb{Z} + a$.

For instance, if we take $n = 2$ and $a = 0$ we are looking at the set

$$\{b : b = 0 + 2q \text{ for some } q \in \mathbb{Z}\} = \{b : b = 2q \text{ for some } q \in \mathbb{Z}\}.$$

which of course is just the set of even integers,

$$\{\ldots, -4, -2, 0, 2, 4, \ldots\},$$

and if in this situation we take instead $a = 1$, we have the set of odd integers

$$\{\ldots, -5, -3, -1, 1, 3, 5, \ldots\}.$$

For $3\mathbb{Z} + 1$ is the set

$$[\, 1 \,]_3 = \{\ldots, -5, -2, 1, 4, 7, 10, \ldots\}.$$

We should point out that (in accordance with Theorem 2.82) we could just have easily denoted this set as $[\,a\,]_3$ where a is any element in $[\,1\,]_3$. For instance we have that

$$[\,-5\,]_3 = [\,-2\,]_3 = [\,1\,]_3 = [\,4\,]_3 = [\,7\,]_3 = [\,10\,]_3\,.$$

By Corollary 3.30 there exists a unique positive integer $r \in [\,a\,]_n = n\mathbb{Z} + a$ so that $0 \leq r < n$. Indeed, this r is the remainder of dividing n by a. Thus, since we have an equivalence relation, we can rephrase this to say that there exists a unique positive integer $0 \leq r < n$ so that

$$[\,r\,]_n = [\,a\,]_n\,.$$

In summary,

THEOREM 3.120. *For every class $[\,a\,]_n$ there exists a unique integer r so that*

(1) $[\,r\,]_n = [\,a\,]_n$,

(2) $0 \leq r < n$.

Moreover this integer r is the remainder of dividing n by a, i.e. $r_{(n)}(a)$ from Notation 3.28. ■

DEFINITION 3.121. For any integer $a \in \mathbb{Z}$, a mod n is defined to be the unique integer r, $0 \leq r < n$, with $[\,a\,]_n = [\,r\,]_n$ which is guaranteed by the theorem.

Note that

$$\begin{aligned} \mathbb{Z} &\to \{0, \ldots, n-1\} \\ a &\mapsto (a \bmod n) \end{aligned}$$

is a function: the existence part guarantees that $a \mapsto a$ mod n is defined everywhere, and the uniqueness part guarantees that it is single valued. This function is surjective since if $0 \leq r < n$, then r mod $n = r$. Note also that if $r \in \{0, \ldots, n-1\}$, then r mod $n = r$ and if $r \neq s \in \{0, \ldots, n-1\}$, then $[\,r\,]_n \neq [\,s\,]_n$ since r mod $n = r \neq s = s$ mod n. Again, the way to find this integer is by finding the remainder of a after dividing n into it.

In fact we have

PROPOSITION 3.122. *For any two integers n and a*

$$r_{(n)}(a) = a \bmod n.$$

■

WARNING! 3.123. Often textbooks call the relation itself mod n. We do not do this. Therefore, to us we have, for instance,

$$9 \bmod 7 = 2 \quad \text{but} \quad 2 \bmod 7 \neq 9,$$

indeed it equals 2.

THEOREM 3.124. TFAE

(1) $n \mid (a - b)$

(2) $a \equiv_n b$,

(3) $a \bmod n = b \bmod n$,

(4) $[\, a \,]_n = [\, b \,]_n$,

(5) $b \in n\mathbb{Z} + a$,

(6) $r_{(n)}(a) = r_{(n)}(b)$ [II]

Proof. Most of this is just a restatement of the fact that $a \equiv_n b$ is an equivalence relation. We only prove that $a \equiv_n b$ holds iff $a \bmod n = b \bmod n$. Let r denote the integer $(a \bmod n)$ and s denote the integer $(b \bmod n)$.

$\Rightarrow$ If $a \equiv_n b \Rightarrow [\, a \,]_n = [\, b \,]_n$. By definition, $[\, a \,]_n = [\, r \,]_n$ and $[\, b \,]_n = [\, s \,]_n$. Therefore, $[\, r \,]_n = [\, s \,]_n$, which implies that $r = s$.

$\Leftarrow$ Suppose that $r = s$. Then,

$$[\, a \,]_n = [\, r \,]_n = [\, s \,]_n = [\, b \,]_n \Rightarrow a \equiv_n b$$

as desired. ■

Corollary 3.125.

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

with no repeats. In particular, $\#(\mathbb{Z}_n) = n$.

Proof. Let $\underline{n} = \{0, \ldots, n-1\}$. For any $i, j \in \underline{n}$ we have that $i \bmod n = j \bmod n$ holds iff $i = j$ iff $[\, i \,]_n = [\, j \,]_n$. Hence the map

$$\begin{aligned} \pi_n \colon \{0, \ldots, n-1\} &\to \mathbb{Z}_n \\ : i &\mapsto [\, i \,]_n \end{aligned}$$

is an injection. Since every class $[\, a \,]_n$ is equal to some $[\, i \,]_n$ for some $i \in \underline{n}$ we also have that it's surjective. Hence a bijection. ■

Remark 3.126. A fancier way of stating this is that the function

$$\begin{aligned} \bmod n \colon \mathbb{Z} &\to \{0, \ldots, n-1\} \\ : m &\mapsto (m \bmod n) \end{aligned}$$

induces a well–defined function

$$\begin{aligned} \widetilde{\bmod}\, n \colon \mathbb{Z}_n &\to \{0, \ldots, n-1\} \\ : [\, m \,]_n &\mapsto (m \bmod n) \end{aligned}$$

which is a bijection.

[II] In words, dividing n by a has the same remainder as dividing n by b.

Example 3.127.

(1) We have that $12315 \equiv_7 2$ and $12315 \equiv_7 9$, but we only have

$$(12315 \bmod 7 = 2) \quad \textbf{but} \quad (12315 \bmod 7 \neq 9).$$

since $9 > 7$. Note that $9 \bmod 7 = 2$.

(2) The integer $123162 = 784 \cdot 157 + 74$. Therefore, $123162 \bmod 157 = 74$. For another example consider

$$78260 = 498 \cdot 157 + 74$$

which implies that $78260 \bmod 157 = 74$ as well. Using the notation correctly we have

$$[\,123162\,]_{157} = [\,74\,]_{157} = [\,78260\,]_{157}\,.$$

The corollary means that every integer is equal to a unique integer mod$n$. In the language of equivalence relations, this means the numbers $\ln = \{0, 1, \ldots, n-1\}$ make up a set of representatives for the partition defined by the equivalence relation. Just like every rational number is typically reduced so that the numerator and denominator are relatively prime to each other, we typically write $[\,a\,]_n$ as the integer it is equal to mod$n$. However, again in analogy with the rational numbers, we will naturally need to "leave" the reduced world to be able to actually do computations. For instance, in $\mathbb{Q}$ in order to add say $1/6$ and $2/15$ we first convert both fractions to the unreduced forms, $5/30$ and $4/30$, then we add them to get $9/30$. Lastly we reduce again to get a final answer of $3/10$.

This also allows us to be sloppy from time to time and it allows us to "forget" that we are dealing with classes. However, this sloppiness often allows students to make the following huge mistake:

Warning! 3.128. The set $\mathbb{Z}_n$ is **NOT** the literally the same as the set of integers $0, \ldots, n-1$. In fact elements of $\mathbb{Z}_n$ are not integers at all, they are classes, each an infinite set of integers. The integers $0, \ldots, n-1$ simply make logical choices of *representatives* of these classes. Remember, that for instance $[\,1\,]_2$ is the set of all odd integers (and the element 1 happens to be one such odd integer)! Calling an element of $\mathbb{Z}_n$ an integer is like calling Dick Durbin a state ... simply because he represents one (Illinois).[III]

As a teacher, this is difficult because my goal is for you to be able to walk the fine line of dealing with the classes in $\mathbb{Z}_n$ as if they were integers (and not obsessing about the fact that they are classes) and making you keep in the back of your mind that these dudes really are classes. This is really easy for us in the case of the rational numbers and hopefully, with a little bit of practice, should become easy as well in this case. Perhaps the best analogy is with the days of the week. Instead of calling the elements of $\mathbb{Z}_7$, $[\,0\,]_5\,, [\,1\,]_5\,, \ldots, [\,6\,]_5$ perhaps you should call them

$$\{\text{Sunday}, \text{Monday}, \ldots, \text{Saturday}\}$$

(in that order) for a bit to really drive this into your head.

We will attempt to stick with the following convention.

Convention 3.129. Unless otherwise stated, when writing down a class for an element of $\mathbb{Z}_n$ we assume we writing it in terms of the representative in $\{0, \ldots, n-1\}$.

Again, compare this convention with what we do for rational numbers where we always agree to write a rational number in "reduced" form of a/b where a and b are relatively prime. Of course it wouldn't be wrong to say that we have $9/12$ of a pizza left, this is not done, and we say that we have $2/3$ of a pizza left. Same thing with integers mod n. While it's not wrong to write down an answer like $[\,17\,]_5$, so everybody knows we're talking about the same thing we'll stick to writing this as $[\,2\,]_5$.

3.4.2 Operations in $\mathbb{Z}_n$

The main thing which makes the sets $\mathbb{Z}_n$ so useful is that classes can be added, subtracted, multiplied, and (often) divided (for a fixed n). For instance, if we consider the classes

$$[\,2\,]_7 = \{\ldots, -12, -5, 2, 9, 16, \ldots\} \quad \text{and} \quad [\,4\,]_7 = \{\ldots, -10, -3, 4, 11, 18, \ldots\}$$

then for any element of the first set, $z_1 \in [\,2\,]_7$, and any element of the second set $z_2 \in [\,4\,]_7$ we can hopefully convince ourselves that $z_1 + z_2 \in [\,6\,]_7$. For instance,

$$(-12) + (-3) = -15 \equiv_7 -15 + 21 = 6.$$

To actually prove this note that such a z_1 would equal $2 + 7m$ for some integer $m \in \mathbb{Z}$ and z_2 would equal $4 + 7m'$ for some integer $m' \in \mathbb{Z}$. Therefore,

$$z_1 + z_2 = (2 + 7m) + (4 + 7m') = 6 + 7(m + m') = 6 + 7n$$

for some integer n (namely $n = m + m'$). Not only that, but if we instead multiply any such z_1 and z_2 we get that

$$z_1 \cdot z_2 = (2 + 7m) \cdot (4 + 7m') = 2 \cdot 4 + 2 \cdot 7m' + 4 \cdot 7m + 7m \cdot 7m' = 8 + 7n' = 1 + 7n''$$

for some integers n' and n'' (namely $n' = 2 + 4 + 7 = 13$ and $n'' = n' + 1$). This proves that if $z_1 \in [\,2\,]_7, z_2 \in [\,4\,]_7$ then

$$z_1 + z_2 \in [\,6\,]_7 \quad \text{and} \quad z_1 \cdot z_2 \in [\,1\,]_7\,.$$

If we replace the specific integers, 2, 4, and 7 by variables we can prove that

Theorem 3.130. *For any positive integer n, multiplication and addition mod n are well defined. That is*

$$a \equiv_n a',\ b \equiv_n b', \Rightarrow ab \equiv_n a'b',\ a + b \equiv_n a' + b'.$$

Proof. It's a bit more difficult to prove this for multiplication, so we only do that case. The assumption guarantees that there are integers m and m' so that $a = a' + nm$ and $b = b' + nm'$. Therefore

$$a \cdot b = (a' + nm)(b' + nm') = a'b' + b'nm + a'nm' + nnmm' = a' \cdot b' + nr$$

for some integer r (namely $r = b'm + a'm' + nmm'$). Therefore $a \cdot b - a' \cdot b'$ is divisible by n and as such $a \cdot b \equiv_n a' \cdot b'$ as desired. ∎

This means the following definition is well defined,

Definition 3.131. Let $[\,a\,]_n, [\,b\,]_n \in \mathbb{Z}_n$. Then we define

$$[\,a\,]_n \oplus [\,b\,]_n \stackrel{\text{def}}{=} [\,a + b\,]_n \quad \text{and} \quad [\,a\,]_n \odot [\,b\,]_n \stackrel{\text{def}}{=} [\,a \cdot b\,]_n\,.$$

As we have done several times already, we drop the circles around our operations. Therefore we have the definitions

$$[\,a\,]_n + [\,b\,]_n \stackrel{\text{def}}{=} [\,a + b\,]_n \quad \text{and} \quad [\,a\,]_n \cdot [\,b\,]_n \stackrel{\text{def}}{=} [\,a \cdot b\,]_n\,.$$

Notice that before just a moment ago we have no definition for $[\,a\,]_n + [\,b\,]_n$ or $[\,a\,]_n \cdot [\,b\,]_n$.

Definition 3.132. We for $m > 0$, define the m^{th} *power* of a class $[\,a\,]_n$ as the product with itself m times:

$$([\,a\,]_n)^m \overset{\text{def}}{=} \underbrace{[\,a\,]_n \cdot \cdots \cdot [\,a\,]_n}_{n\text{–times}}.$$

We will make implicit use of the following exercise.

Exercise 3.52. Prove that $([\,a\,]_n)^m = [\,a^m\,]_n$.
Careful! Pay close attention to what both sides mean.

Example 3.133. The following tables depict the operations in $\mathbb{Z}_5$.

$+$	$[0]_5$	$[1]_5$	$[2]_5$	$[3]_5$	$[4]_5$
$[0]_5$	$[0]_5$	$[1]_5$	$[2]_5$	$[3]_5$	$[4]_5$
$[1]_5$	$[1]_5$	$[2]_5$	$[3]_5$	$[4]_5$	$[0]_5$
$[2]_5$	$[2]_5$	$[3]_5$	$[4]_5$	$[0]_5$	$[1]_5$
$[3]_5$	$[3]_5$	$[4]_5$	$[0]_5$	$[1]_5$	$[2]_5$
$[4]_5$	$[4]_5$	$[0]_5$	$[1]_5$	$[2]_5$	$[3]_5$

$\cdot$	$[0]_5$	$[1]_5$	$[2]_5$	$[3]_5$	$[4]_5$
$[0]_5$	$[0]_5$	$[0]_5$	$[0]_5$	$[0]_5$	$[0]_5$
$[1]_5$	$[0]_5$	$[1]_5$	$[2]_5$	$[3]_5$	$[4]_5$
$[2]_5$	$[0]_5$	$[2]_5$	$[4]_5$	$[1]_5$	$[3]_5$
$[3]_5$	$[0]_5$	$[3]_5$	$[1]_5$	$[4]_5$	$[2]_5$
$[4]_5$	$[0]_5$	$[4]_5$	$[3]_5$	$[2]_5$	$[1]_5$

We can use the well defined nature of these operations to our advantage by picking representatives for the class in question which makes it easiest to perform the actual operation in question. For instance, to calculate what $1123 \cdot 1125$ is mod 1126 its much easier to use the representatives

$$-3 \in [\,1123\,]_{1126} \quad \text{and} \quad -1 \in [\,1125\,]_{1126}$$

to see right away that this product is equivalent to 3 mod 1,126. In more detail,

$$[\,1123\,]_{1126} \cdot [\,1125\,]_{1126} = [\,-3\,]_{1126} \cdot [\,-1\,]_{1126} = [\,(-3)\cdot(-1)\,]_{1126} = [\,3\,]_{1126}.$$

Indeed, if we did this out the long way we would first need to do long multiplication to figure out that

$$1123 \cdot 1125 = 1,263,375$$

Then we would have to do a long division to see that this number, $1,263,375$ when divided by $1,126$ gives

$$1,263,375 = (1122)\cdot(1126) + 3^{\text{IV}}$$

implying that the product is equal to 3 mod 1126, which of course matches what we got doing this the easy way.

[IV] Actually, I cheated here and used my calculator to figure out what $1,263,375 - 3 = 1,263,372$ is when divided by $1,126$.

Another cool trick is with powers. For instance, note that $[\,5^2\,]_{13} = [\,-1\,]_{13}$ and $[\,5^4\,]_{13} = [\,1\,]_{13}$. As such, we have

$$\begin{aligned}
[\,5\,]_{13}^{1033} &= [\,5\,]_{13}^{1032+1} \\
&= [\,5\,]_{13}^{4\cdot 258+1} \\
&= [\,5\,]_{13}^{4\cdot 258}\,[\,5^1\,]_{13} \\
&= \left([\,5\,]_{13}^{4}\right)^{258} \cdot [\,5\,]_{13} \\
&= [\,1\,]_{13}^{258} \cdot [\,5\,]_{13} \\
&= [\,1\,]_{13} \cdot [\,5\,]_{13} \\
&= [\,5\,]_{13}\,.
\end{aligned}$$

Actually, the same basic proof would show that

$$\begin{aligned}
a \equiv_4 0 &\Rightarrow [\,5\,]_{13}^{a} = [\,1\,]_{13} \\
a \equiv_4 1 &\Rightarrow [\,5\,]_{13}^{a} = [\,5\,]_{13} \\
a \equiv_4 2 &\Rightarrow [\,5\,]_{13}^{a} = [\,-1\,]_{13} = [\,12\,]_{13} \\
a \equiv_4 3 &\Rightarrow [\,5\,]_{13}^{a} = [\,-5\,]_{13} = [\,8\,]_{13}\,.
\end{aligned}$$

EXERCISE 3.53. Calculate

(1) $[\,16\,]_{18} + [\,11\,]_{18}$

(2) $[\,6\,]_{18} \cdot [\,3\,]_{18}$

(3) $[\,1012\,]_{1015} \cdot [\,1012\,]_{1015}$

(4) $[\,2^{1035}\,]_{7}$

(5) $[\,3^{1035}\,]_{11}$

(All answers should be given in terms of integers between 0 and $n-1$).

For every n, the properties of these operations are quite similar to those for $\mathbb{C}$.

THEOREM 3.134. *Fix a positive integer n. The following hold*

(1) (Identity Properties)

$$\begin{aligned}
[\,0\,]_n + [\,a\,]_n &= [\,a\,]_n = [\,a\,]_n + [\,0\,]_n \\
[\,1\,]_n \cdot [\,a\,]_n &= [\,a\,]_n = [\,a\,]_n \cdot [\,1\,]_n
\end{aligned}$$

(2) (Inverse Properties)

$$\exists\ [\,x\,]_n \in \mathbb{Z}_n \ s.t\ [\,a\,]_n + [\,x\,]_n = [\,0\,]_n = [\,x\,]_n + [\,a\,]_n\,.$$

(3) (Distributive)

$$[a]_n \cdot ([b_1]_n + [b_2]_n) = [a]_n \cdot [b_1]_n + [a]_n \cdot [b_2]_n$$
$$([a_1]_n + [a_2]_n) \cdot [b]_n = [a_1]_n \cdot [b]_n + [a_2]_n \cdot [b]_n .$$

(4) (Associative Properties)

$$([a_1]_n \cdot [a_2]_n) \cdot [a_3]_n = [a_1]_n \cdot ([a_2]_n \cdot [a_3]_n)$$
$$([a_1]_n + [a_2]_n) + [a_3]_n = [a_1]_n + ([a_2]_n + [a_3]_n)$$

(5) (Commutative Properties)

$$[a]_n + [b]_n = [b]_n + [a]_n$$
$$[a]_n \cdot [b]_n = [b]_n \cdot [a]_n .$$

Proof. These are all quite easy to prove once one understands the basic trick involved. Roughly, we take the corresponding identities for the integers and take them mod n. I illustrate this idea only once; by proving that multiplication is associative:

$$\begin{aligned}
([a_1]_n \cdot [a_2]_n) \cdot [a_3]_n &= [(a_1 \cdot a_2)]_n \cdot [a_3]_n && \text{(Def of mult mod } n\text{)} \\
&= [(a_1 \cdot a_2) \cdot a_3]_n && \text{(Def of mult mod } n\text{)} \\
&= [a_1 \cdot (a_2 \cdot a_3)]_n && \text{(Mult in } \mathbb{Z} \text{ is assoc.)} \\
&= [a_1]_n \cdot [a_2 \cdot a_3]_n && \text{(Def of mult mod } n\text{)} \\
&= [a_1]_n \cdot ([a_2]_n \cdot [a_3]_n). && \text{(Def of mult mod } n\text{)}
\end{aligned}$$

All other properties that were listed are proved in the same way: combine everything inside of a bracket, then use the corresponding property in $\mathbb{Z}$, and lastly decompose everything back out. ■

Definition 3.135. Let R be a ring[V] (see definition at the beginning of the chapter). We say that an element $r \in R$ is invertible if there exists an element $s \in R$ so that $sr = rs = 1$. The set of invertible elements is denoted as $\mathcal{U}(R)$. Such an element is called an (can be shown to be unique!) inverse for r.

Warning! 3.136. In terms of the future notation from the next chapter, $\mathcal{U}(R)$ means $\mathcal{U}(R, \cdot)$.

Definition 3.137. A non-zero element r in a ring R is called a *zero divisor* if there exists a non-zero element $s \in R$ so that $rs = 0$.

In symbols, if $r \neq 0$ is *not* a zero divisor then we have

$$r\,s = 0 \Rightarrow s = 0.$$

Exercise 3.54. Show that a ring R contains no zero-divisors iff $R^\times$ is closed under multiplication.
Hint! This is essentially just a translation of definitions.

SOLUTION

For the rings $\mathbb{H}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ every non-zero element is invertible, i.e. $\mathcal{U}(\mathbb{H}) = \mathbb{H}^\times$, etc. For the ring of integers, only 1 and -1 are invertible.[VI] That is $\mathcal{U}(\mathbb{Z}) = \{\pm 1\}$. In all of these rings, zero divisors do not exist. The question of which elements of $\mathbb{Z}_n$ are invertible and which are zero divisors is more exciting. For instance the elements $[\,2\,]_6$ and $[\,3\,]_6$ are zero divisors (in $\mathbb{Z}_6$) since

$$[\,2\,]_6 \cdot [\,3\,]_6 = [\,6\,]_6 = [\,0\,]_6 = 0.$$

Theorem 3.138. *If an element of a ring is invertible, then it is not a zero divisor.*

Proof. Suppose that r is invertible with inverse r' and that $r\,s = 0$. We need to prove that $s = 0$. Then

$$0 = r' \cdot 0 = r' \cdot (r \cdot s) = (r' \cdot r) \cdot s = 1 \cdot s = s.$$

Therefore $s = 0$ and r is not a zero divisor. ∎

Theorem 3.139. *Let n and m be integers. Then* TFAE

(1) $\gcd(n, m) = 1$
(2) $[\,m\,]_n \in \mathcal{U}(\mathbb{Z}_n)$[VII]
(3) $[\,m\,]_n$ *is not a zero divisor in* $\mathbb{Z}_n$.

In words $[\,m\,]_n$ is invertible in $\mathbb{Z}_n$ iff m and n are relatively prime iff it is not a zero divisor.

Proof. This theorem is logically equivalent to the Euclidean algorithm formula given in Theorem 3.31. The proof illustrates this.

1 ⇒ 2 Suppose that n and m are relatively prime. Then by the Euclidean algorithm formula, 3.31, we have integers a and b so that $an + bm = 1$. But this means that

$$\begin{aligned}
&[\,an + bm\,]_n = [\,1\,]_n \\
\Rightarrow\ &[\,an\,]_n + [\,bm\,]_n = [\,1\,]_n \\
\Rightarrow\ &[\,a\,]_n\,[\,n\,]_n + [\,b\,]_n\,[\,m\,]_n = [\,1\,]_n \\
\Rightarrow\ &[\,a\,]_n\,[\,0\,]_n + [\,b\,]_n\,[\,m\,]_n = [\,1\,]_n \\
\Rightarrow\ &[\,0\,]_n + [\,b\,]_n\,[\,m\,]_n = [\,1\,]_n \\
\Rightarrow\ &[\,b\,]_n \cdot [\,m\,]_n = [\,1\,]_n .
\end{aligned}$$

This proves that $[\,b\,]_n$ is an (multiplicative) inverse for $[\,m\,]_n$. In particular, one exists.

2 ⇒ 3 This was the lemma.

2 ⇒ 1 Suppose that m has a mod n inverse, say b. Then we have that $bm \equiv_n 1$ which means that $n \mid (bm - 1)$. But this means that there exists an integer a' so that $a'n = bm - 1$. Rearranging shows that $1 = bm + an$ (with $a = -a'$). But this guarantees that m and n are relatively prime.

[VI] While you might be tempted to say that a number like 2 is invertible (with inverse 1/2) this would be wrong since we require that the inverse be in the same ring! Since 1/2 is **not** an integer it can't qualify for an inverse in this context!
[VII] Again, this means there exists an integer a so that $am \equiv_n 1$.

3 ⇒ 1 We prove the contrapositive. Suppose that $d \mid m$ and $d \mid n$ with $d > 1$ (i.e we are assuming not **(1)**). Since $d \mid m$ and $d \mid n$ there exist positive integers $a < m$ and $b < n$ so that

$$m = ad \quad \text{and} \quad n = bd.$$

Therefore

$$bm = bad = a(bd) = an \Rightarrow [\, b \,]_n [\, m \,]_n = [\, a \,]_n [\, n \,]_n = [\, 0 \,]_n .$$

But since $b < n$ we have that $[\, b \,]_n \neq [\, 0 \,]_n = 0$. This implies that $[\, m \,]_n$ is a zero divisor (therefore we have proved not **(3)**). ■

In symbols, this theorem states that

COROLLARY 3.140.

$$\mathcal{U}(\mathbb{Z}_n) = \{[\, a \,]_n : \gcd(a, n) = 1\}$$

In words, the following Corollary says that every non-zero element of $\mathbb{Z}_p$ is invertible. Recall that for a ring R, the notation $R^\times$ means $R \setminus \{0\}$, i.e. all non-zero elements of R.

COROLLARY 3.141. *If p is prime, then*

$$\mathcal{U}(\mathbb{Z}_p) = \mathbb{Z}_p^\times.$$

DEFINITION 3.142. A ring R with the property that $\mathcal{U}(R) = R^\times$ is called a *skew-field*. A commutative skew field is called a P *field*.

Therefore, using the fact that $\mathbb{Z}_p$ is commutative, the corollary just above, Corollary 3.141 says that $\mathbb{Z}_p$ is a field. Other fields include $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$. $\mathbb{H}$ is a skew field (but is not a field). The integers, $\mathbb{Z}$, and $\mathbb{Z}_n$, for n a composite number, are not skew-fields. The definitions are not really a big part of the main part of the course, but they are among the most important definitions in an abstract algebra course which focuses on rings. We cover them in more detail in Appendix E. If we want to stick to the sloppyness of Temporary Definition 3.1 which we started the chapter with we have

TEMPORARY DEFINITION 3.143. A field is one of the sets $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$, or one of the sets $\mathbb{Z}_p$ for p a prime.

Of course there are many more fields than just this basic examples. For instance, once we generalize the definition given at the beginning of the chapter of a ring (we do this in the appendix), we will see that

$$\mathbb{Q}[i] \stackrel{\text{def}}{=} \{a + bi : a, b \in \mathbb{Q}\}$$

forms a ring which is moreover a field which lies properly between $\mathbb{Q}$ and $\mathbb{R}$ (since $\mathbb{Q}[i]$ contains i but it does not contain $\sqrt{2}$.)

PROOF. For any $[\, a \,]_p \in \mathbb{Z}_p$ either $p \mid a$ in which case $[\, a \,]_p = [\, 0 \,]_p$ or $\gcd(p, a) = 1$ in which case $[\, a \,]_p \in \mathcal{U}(\mathbb{Z}_p)$ by Corollary 3.140. ■

REMARK 3.144. Of course there are an infinite number of integers relatively prime to n, but there are only a finite number of classes relatively prime to n. A Corollary of the above could be that if $a \equiv_n b$, then a is relatively prime to n iff b is.

Example 3.145. The theorem guarantees us that

$$\mathcal{U}(\mathbb{Z}_7) = \{[1]_7, [2]_7, [3]_7, [4]_7, [5]_7, [6]_7\} = \mathbb{Z}_7^\times$$
$$\mathcal{U}(\mathbb{Z}_9) = \{[1]_9, [3]_9, [7]_9, [9]_9\}$$
$$\mathcal{U}(\mathbb{Z}_{30}) = \{[1]_{30}, [7]_{30}, [11]_{30}, [13]_{30}, [17]_{30}, [19]_{30}, [23]_{30}, [29]_{30}\}$$
$$\mathcal{U}(\mathbb{Z}_{36}) = \{[1]_{36}, [5]_{36}, [7]_{36}, [11]_{36}, [13]_{36}, [17]_{36}, [19]_{36}, [23]_{36}, [25]_{36}, [29]_{36}, [31]_{36}, [35]_{36}\}.$$

For instance we have that $[7]_{36}^{-1} = [31]_{36}$. This is because

$$[7]_{36} \cdot [31]_{36} = [1]_{36}$$

as can easily be checked (easiest way to do this quickly is change the class of 31 to the class of -5).

3.4.3 Euler Totient Function

Definition 3.146. For any positive integer n, the *Euler totient* function, ϑ is defined in any of the following equivalent ways

$$\begin{aligned}\vartheta(n) &\overset{\text{def}}{=} \text{Number of integers } i \text{ s.t. } 1 \leq i \leq n \text{ and } \gcd(i,n) = 1\\ &= \#\{i \in \overline{n} : \gcd(i,n) = 1\}\\ &= \#\mathcal{U}(\mathbb{Z}_n).\end{aligned}$$

SubDefinition 3.147. For an integer m we define the *m reduced Euler totient function* as

$$\vartheta_m(n) \overset{\text{def}}{=} \begin{cases}\vartheta(n) & \text{if } n \mid m\\ 0 & \text{otherwise}\end{cases}.$$

Note that since 1 is only divisible by itself, we have that $\vartheta_1 = 0$.
We will show that we have the numerical formula

$$\vartheta(n) = n \cdot \prod_{\substack{p \mid n\\ p>0 \text{ prime}}} \left(1 - \frac{1}{p}\right). \tag{40}$$

[VIII]

We give half of the proof now and later in Corollary 8.78 coming up on page 428. (We will also give a slick proof as a Corollary of some general group theory in 5.96 coming up on page 297 in Chapter 5.) (So p ranges over the *distinct* primes dividing n.) Perhaps the most elementary way of showing this is by an inclusion, exclusion argument. We start a different proof which minimizes detours.

First off, for powers of primes p^r, we show that

$$\vartheta(p^r) = p^r \cdot \left(1 - \frac{1}{p}\right) = p^{r-1} \cdot (p-1).$$

Indeed, this is easy to prove directly since the only numbers between 1 and p^r which are not relatively prime to p^r are the p^{r-1} multiples of p:

$$p,\ 2p,\ 3p,\ \ldots,\ p^{r-1} \cdot p = p^r.$$

Hence

$$\vartheta(p^r) = p^r - p^{r-1} = p^{r-1}(p-1)$$

as stated. Using this for powers of a prime, and the fact, proven later in Corollary 8.78 coming up on page 428, for numbers n and m which are relatively prime we have that

$$\vartheta(nm) = \vartheta(n) \cdot \vartheta(m)$$

we can recover the general formula by factoring n into a product of distinct primes, $n = p_1^{r_1} \cdots p_m^{r_m}$ and

$$\begin{aligned}
\vartheta(p_1^{r_1} \cdots p_m^{r_m}) &= \vartheta(p_1^{r_1}) \cdots \vartheta(p_m^{r_m}) \\
&= p_1^{r_1} \cdot \left(1 - \frac{1}{p_1}\right) \cdots p_m^{r_m} \cdot \left(1 - \frac{1}{p_m}\right) \\
&= \underbrace{p_1^{r_1} \cdots p_m^{r_m}}_{n} \cdot \left[\left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_m}\right)\right] \\
&= n \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right).
\end{aligned}$$

For instance

$$\vartheta(30) = 30 \cdot \left(1 - \frac{1}{2}\right) \cdot \left(1 - \frac{1}{3}\right) \cdot \left(1 - \frac{1}{5}\right) = 30 \cdot \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} = 8$$

which matches what we calculated above in Example 3.145 (since we found 8 elements in $\mathcal{U}(\mathbb{Z}_{30})$).

For another example, using a computer program or playing with the numbers (for longer than I'd care to!) one can show that the only primes which divide the number 492, 156 are 2, 3, 7, and 31. [IX] Therefore

$$\vartheta(492,156) = 492,156 \cdot \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right) \cdot \left(1 - \frac{1}{7}\right) \cdot \left(1 - \frac{1}{31}\right) = \frac{492156 \cdot 2 \cdot 6 \cdot 30}{2 \cdot 3 \cdot 7 \cdot 31} = 136,080.$$

You can find more information about this cool function at wiki

http://en.wikipedia.org/wiki/Euler's totient function

and a handy java calculator (for small values of n) at

http://professorjava.weebly.com/totient.html.

If you wish to calculate larger values you can turn to `MAPLE` by using the following commands:

[IX] Computationally, finding which primes divide a number is clearly only slightly easier than finding the entire factorization of that number. However, for us as humans, we'll see that this will be considerably easier! Of course it's easy to check that the primes listed divide 492, 156 but how would you know that if you weren't told it?

```
> with(numtheory):
> phi(229320);
```

Exercise 3.55. Assuming (40) to be valid prove that

$$\vartheta_m(n) = n \cdot \prod_{\substack{p|n \\ p>0 \text{ prime}}} \left(1 - \frac{1}{\gcd(m,p)}\right)$$

Exercise 3.56. List all elements in $\mathcal{U}(\mathbb{Z}_{18})$ and pair each one with their inverse.

Exercise 3.57. Directly[x] find the inverses of the following or state that an inverse does not exist:

(1) $[5]_{21}$ **(2)** $[4]_{17}$ **(3)** $[7]_{11}$ **(4)** $[21]_{35}$.

SOLUTION

3.4.4 Computing the Inverse*

It should come as a bit of a shock that, without a stroke of luck, it's really quite difficult to compute the (multiplicative) inverse of a class. This is another aspect which makes $\mathbb{Z}_n$ different from the quaternions, $\mathbb{H}$. In the latter case you should recall how easy it was to write down an explicit formula for the (multiplicative) inverse of an quaternion (see Equation (**??**)), but no such explicit formula exists in our present situation. The best we will be able to do is to come up with an algorithm which takes several steps.

Maple Example 3.148. In MAPLE you can experiment with this for some simple cases by, for example using the command

```
> 35^(-1) mod 83;
```

gives the correct output.

There is an algorithm for computing the inverse of an element mod b (assuming the arguments are relatively prime so that one exists). Given relatively prime integers a and b which after possibly replacing a with a mod b we assume that $a < b$. Then

Algorithm 3.149.

Step 1 Compute all quotients by applying the division algorithm (see Remark **??**). That is define $r_{-1} = b$, $r_0 = a$ and

$$\begin{aligned} b &= q_0 a + r_1 \\ a &= q_1 r_1 + r_2 \\ r_1 &= q_2 r_2 + r_3 \\ &\vdots \quad \vdots \\ r_{n-1} &= q_n r_n + 1 \end{aligned}$$

[x] Meaning do not use the methods from the next section, Section 3.4.4.

Step 2 Set up a table with two rows and $n+3$ columns in the first row placed over a centered $n+1$ columns in the second. In the 1,1 entry place the number 0 and in the 1,2 entry place the number 1.

$$\begin{array}{cccccccc} 0 & 1 & - & - & - & \cdots & - & \star \\ & - & - & - & - & \cdots & - & \end{array}$$

Step 3 Starting from the 2,2 entry fill in all the q values

$$\begin{array}{cccccccc} 0 & 1 & - & - & - & \cdots & - & \star \\ & q_0 & q_1 & q_2 & q_3 & \cdots & q_n & \end{array}$$

Step 4 Now we fill in the top row of the table inductively by taking the entry two to the left of the one we wish to fill in and subtracting the product of the two entries in the column to the immediate left of the entry we are trying to fill in. That is

$$\begin{array}{ccccccccccccccc} 0 & 1 & x_1 & x_2 & x_3 & \cdots & x_{i-1} & x_i & x_{i+1} := x_{i-1} - x_i q_i & - & \cdots & - & \star \\ & q_0 & q_1 & q_2 & q_3 & \cdots & q_{i-1} & q_i & q_{i+1} & q_{i+2} & \cdots & q_n & \end{array}$$

Step 5 The entry which replaces the $\star$ is the inverse of a mod b.

Example 3.150. Consider $b = 137$ and $a = 9$. Then

$$137 = 15 \cdot 9 + 2 \qquad 9 = 4 \cdot 2 + 1.$$

Therefore $q_0 = 15$ and $q_1 = 4$. So our table progresses as follows

$$\begin{array}{cccc} 0 & 1 & - & \star \\ & 15 & 2 & \end{array} \Rightarrow \begin{array}{cccc} 0 & 1 & -15 & 61 \\ & 15 & 4 & \end{array}.$$

Therefore $[\,9\,]_{137}^{-1} = [\,61\,]_{137}$. Indeed this is the case since

$$9 \cdot 61 = 549 = 4 \cdot 137 + 1 \Rightarrow [\,9\,]_{137} \cdot [\,61\,]_{137} = [\,1\,]_{137}.$$

Actually, this shows us that

$$[\,137\,]_9^{-1} = [\,-4\,]_9 = [\,5\,]_9$$

since the number 4 shows up highlighted in red above. (Why is this? check for yourselves!) Of course this can very easily have been checked by hand since the number 9 is itty–bitty and

$$137 \equiv_9 1 + 3 + 7 = 11 \equiv_9 2.$$

Example 3.151. Let $a = 20$ and $b = 37$. Then

$$37 = 1 \cdot 20 + 17 \quad 20 = 1 \cdot 17 + 3 \quad 17 = 5 \cdot 3 + 2 \quad 5 = 2 \cdot 2 + 1$$

and so $q_0 = 1$, $q_1 = 1$, $q_2 = 5$, $q_3 = 2$ and our table is

$$\begin{array}{cccccc} 0 & 1 & - & - & - & \star \\ & 1 & 1 & 5 & 1 & \end{array} \Rightarrow \begin{array}{cccccc} 0 & 1 & -1 & 2 & -11 & 13 \\ & 1 & 1 & 5 & 1 & \end{array}$$

and the inverse of $[\,20\,]_{37}$ is therefore $[\,13\,]_{37}$.

Finding the following inverses without using the algorithm would be more difficult than was the case in Exercise 3.57 back on page 155.

Exercise 3.58. Compute the inverse of the following or state that an inverse does not exist:

(1) $[\,15\,]_{101}$ **(4)** $[\,56\,]_{175}$ **(7)** $[\,35\,]_{359}$

(2) $[\,19\,]_{62}$ **(5)** $[\,41\,]_{83}$ **(8)** $[\,50\,]_{3307}$

(3) $[\,11\,]_{317}$ **(6)** $[\,417\,]_{1984}$ **(9)** $[\,100\,]_{231}$

SOLUTION

3.4.5 Fundamental Theorem of Arithmetic

Corollary 3.152. *Suppose that a and b are relatively prime integers and that $a \mid nb$ for some integer n. Then $a \mid n$.*

Proof. This is just a translation of the fact that invertible elements are not zero divisors already proven. In more detail: the assumption guarantees us that $[\,n\,]_a \cdot [\,b\,]_a = [\,0\,]_a$. But since a and b are relatively prime, there exists a $\widetilde{b} \in \mathbb{Z}_a$ which is an inverse for b. This implies that

$$\begin{aligned}[\,n\,]_a &= [\,n \cdot 1\,]_a \\ &= [\,n\,]_a \cdot [\,1\,]_a \\ &= [\,n\,]_a \cdot [\,b\,]_a \cdot [\,\widetilde{b}\,]_a \\ &= [\,0\,]_a \cdot [\,\widetilde{b}\,]_a \\ &= [\,0\,]_a\,.\end{aligned}$$

But the statement that $[\,n\,]_a = [\,0\,]_a$ is the same as the statement $a \mid n$ as desired. ∎

Corollary 3.153. *For any two integers n and m, we have*

$$nm = \operatorname{lcm}(n,m) \cdot \gcd(n,m).$$

Proof. As usual, to avoid annoyance, we assume that n and m are positive. We leave the more general case, as usual, to you. Let's denote $\gcd(n,m)$ as D. Consider the quantity

$$\frac{nm}{D}.$$

We denote this quantity by $\mathcal{L}$. We need to prove that $\mathcal{L} = \operatorname{lcm}(n,m)$. We do this by proving that $\mathcal{L}$ has the two properties, **lcm1** and **lcm2** (from page 76), needed to be the least common multiple.

lcm1 By definition of the greatest common divisor, both $\frac{m}{D}$ and $\frac{n}{D}$ are integers. Therefore writing $\mathcal{L}$ as

$$n \cdot \left(\frac{m}{D}\right) \quad \text{or} \quad \left(\frac{n}{D}\right) \cdot m$$

we have that $\mathcal{L}$ is equal to both n times an integer and m times an integer. In other words both n and m divide $\mathcal{L}$.

lcm2 Now suppose that some integer N is also divisible by both n and m. Then we have that

$$\begin{aligned} N &= q'n \\ &= q'\frac{n}{D}D \\ \Rightarrow \frac{N}{D} &= q'\frac{n}{D}. \end{aligned} \tag{41}$$

Similarly, there exists an integer q'' so that

$$\begin{aligned} \frac{N}{D} &= q''\frac{m}{D} \\ \Rightarrow & q'\frac{n}{D} = q''\frac{m}{D} \\ \Rightarrow & \frac{m}{D} \Big| q'\frac{n}{D}. \end{aligned}$$

In a previous exercise, Exercise 3.11 back on page 84, you proved that the integers $\frac{n}{D}$ and $\frac{m}{D}$ are relatively prime. Thus by the theorem, we have that

$$\frac{m}{D} \Big| q'.$$

That is there exists an integer q''' so that $q' = q'''\frac{m}{D}$. Plugging this back into (41) we have that

$$\frac{N}{D} = q'\frac{n}{D} = q'''\frac{m}{D}\frac{n}{D}.$$

Lastly, clearing a D from the denominators implies that

$$N = q'''\frac{mn}{D} = q'''\mathcal{L}.$$

In other words $\mathcal{L}$ divides N. Therefore $\mathcal{L}$ satisfies **lcm2**. ■

COROLLARY[2] 3.154. *The least common multiple of any two integers exists.*

PROOF. In the proof we showed that the quantity

$$\frac{nm}{\gcd(n,m)}$$

could be used to define the least common multiple of any two integers. ■

COROLLARY[2] 3.155. *Integer n and m are relatively prime iff* $\operatorname{lcm}(n,m) = nm$.

PROOF. By the previous corollary, we know that

$$\operatorname{lcm}(n,m)\gcd(n,m) = nm$$

and, by definition, $\gcd(n,m) = 1$ iff n and m are relatively prime. ■

COROLLARY 3.156. *If p is a prime number and p | ab, then either p | a or p | b.*

PROOF. This corollary follows directly from the previous one because a number is not relatively prime to a prime number iff the number is divisible by that prime number. ■

Corollary 3.157. *If a prime p divides a product of primes*

$$p_1 \cdots p_m$$

then p is equal to one of the prime numbers listed. (That is there is an i between 1 and m so that $p = p_i$).

Proof. Assume this were false. Define b_i to be the integer $p_i \cdots p_m$. With this notation the number in question is b_1. Since $b_1 = p_1 b_2$ and the assumption is that p divides b_1 we must have that $p \mid p_1$ or $p \mid b_2$. Since p_1 is a prime number which we are assuming is not p we must have that the latter is true. Continuing in this way, we see that $p \mid b_i$ for all $i \leq m$. In particular p must divide b_m. But b_m is the prime number p_m. Hence $p = p_m$, a contradiction. ∎

It is interesting to note that we have yet to use the fundamental theorem of arithmetic which says that every integer can be factored uniquely (up to the order of the factors) into a product of primes. As usual, to avoid annoyance we deal only with positive numbers. Of course for a negative number one simply needs to throw a (-1) in front of the factorization of the corresponding positive number.

Theorem 3.158 (The fundamental theorem of arithmatic)**.** *For any positive integer n there exists a unique sequence of primes*

$$p_1 \leq \cdots \leq p_m \tag{42}$$

so that

$$n = p_1 \cdots p_m. \tag{43}$$

Proof.

Uniqueness Suppose there were two sequences of prime numbers so that n is equal to both

$$p_1 \cdots p_m \quad \text{and} \quad q_1 \cdots q_r. \tag{44}$$

After canceling out any common primes, we can assume that both sides of this equation have no primes in common. By definition (44) implies that $p_1 \mid q_1 \cdots q_r$. The previous corollary guarantees us that $p_1 = q_j$ for some integer j between 1 and r which contradicts our assumption that both sides of (44) have no primes in common.

Existence Let n be the least integer which does not have a factorization as in (43). If n was a prime number, then $m = 1$ and $p_1 = n$ would be a factorization. Therefore we have that n is not a prime number. As such there are positive integers a and b so that $n = ab$. But a and b are integers less than n. Therefore, by assumption they have factorizations

$$a = p_1 \cdots p_r \quad \text{and} \quad b = q_1 \cdots q_{r'}$$

into primes. But then

$$n = ab = (p_1 \cdots p_r)(q_1 \cdots q_{r'}) = p_1 \cdots p_r \cdot q_1 \cdots q_{r'}$$

is a factorization of n into primes. ∎

Rant 3.159. Notice that none of the concepts in the section use this factorization! In elementary school this is the starting point for many definitions, however in abstract algebra this is an ending point. Therefore the name "fundamental theorem of Arithmetic" kinda sucks.[XI] Proofs which use the factorization are typically considered ineloquent and should be avoided! However, if you personally feel more confident finding things like the greatest common divisor between two numbers by first factoring them, go right ahead! However, even if you do this, you should know that we have developed a method (indeed a much faster method) for doing this task without factoring the numbers! These types of distinctions will be more important in ring theory course than our group theory course.

Notice that an algorithm which follows this proof is at least as difficult as an algorithm which determines whether or not a given number is prime.

Example 3.160. The number 50, 184 is equal to

$$2 \cdot 2 \cdot 2 \cdot 3 \cdot 3 \cdot 17 \cdot 41.$$

Of course this is more typically written by grouping together the like primes so that

$$n = p_1^{m_1} \cdots p_1^{m_r}$$

for distinct primes p_i. Notice that we have that ${p_i}^{m_i} \mid n$, but no higher power of p_i divides n.

Notation 3.161. With m_i as above $\nu_{p_i}(n) \stackrel{\text{def}}{=} m_i$. We also take $\nu_p(n) = 0$ if p is a prime which does not divide n.

So we have that $\nu_p(n)$ is equal to the number of times the prime p divides the integer n.

So another way of stating the fundamental theorem of arithmatic is by the formula which holds for any integer n we have the following equality:

$$n = \prod_p p^{\nu_p(n)}.$$

In the product, p ranges over all (positive) prime numbers.

Exercise 3.59. For integers n and m and prime p prove that

(1) $\max(\nu_p(n), \nu_p(m)) \leq \nu_p(n+m) \leq \nu_p(n) + \nu_p(m)$;

(2) $\nu_p(n \cdot m) = \nu_p(n) \cdot \nu_p(m)$.

Solution

3.4.6 Extra Questions

For positive integers n and β, there exists a unique sequence

$$0 \leq a_0, a_1, \ldots, a_n < \beta$$

so that

$$n = \sum_{i=0}^{n} a_i \beta^i.$$

You can find these numbers via

$$a_0 := (n \bmod \beta), \qquad a_1 := \left(\frac{n - a_0}{\beta} \bmod \beta\right) \qquad a_2 := \left(\frac{n - a_0 - a_1}{\beta^2} \bmod \beta\right)$$

and in general

$$a_i := \left(\frac{n - \sum_{j=0}^{i-1} a_j \beta^j}{\beta^i} \bmod \beta \right).$$

In this case, we use the notation

$$n = (a_m.a_{m-1}.\ldots.a_1.a_0)_\beta .$$

Of course, this is just the usual base β way of writing a number. In the case that $\beta = 10$, this is what you've been doing for EVER. For instance, if $n = 1017$ and $\beta = 10$, we have $a_0 = 7$, $a_1 = 1$, $a_2 = 0$, $a_3 = 1$, the digits of that number, and all higher a_i's equal to 0 and of course

$$1017 = (1.0.1.7)_{10} .$$

Remark 3.162. Of course the most common base one encounters is base 10, the *decimal* base. Historians aren't 100% sure why this is! Perhaps it's due to the fact that we have 10 fingers, but this is just a guess.

Example 3.163. Besides base 10, other common bases are base 2, the *binary* base, and base 16, the *hexadecimal* base.[XII] Then the number

$$(11.10.14.13.10.11.1.2)_{16}$$[XIII]

Rant 3.164. Unfortunately our system gives us no reasonable way of thinking of must numbers without thinking of them in base 10. For instance how would you refer to the number "five thousand two hundred and eighty" in conversation without using base 10? At least in text you could use roman numerals (which one could argue [not convincingly in my opinion] is still base 10) or you could write that many tick marks.

is the (decimal) number

$$11 \cdot 16^7 + 10 \cdot 16^6 + 13 \cdot 16^5 + 12 \cdot 16^4 + 10 \cdot 16^3 + 11 \cdot 16^2 + 1 \cdot 16 + 2 = 3,135,023,890.$$

The following exercise proves the famous and familiar result that a number is divisible by 3 (or 9) iff the sum of its digits are divisible by 3 (or 9). For instance, the number $1,023,417$ is divisible by 9 (and 3) since the sum of its digits is 18.

Exercise 3.60. Suppose

$$n = (a_m.a_{m-1}.\ldots.a_1.a_0)_\beta .$$

SOLUTION

[XII] In hexadecimal, the number 10 is typically represented as A, 11 has B, 12 as C, 13 as D, 14 as E, and 15 as F.
[XIII] This notation kind sucks because I'm using base 10 implicitly here. See rant below.

(1) Suppose that $d \mid \beta$. Prove that

$$n \equiv_d a_0.$$

Conclude that $d \mid n$ iff $d \mid a_0$.[XIV]

(2) Suppose that $d \mid (\beta - 1)$. Prove that

$$n \equiv_d \sum_{i=1}^{m} a_i.$$

Conclude that

$$d \Big| n \Leftrightarrow d \Big| \sum_{i=1}^{m} a_i.$$

[XV]

Hint! The assumption is that $b \equiv_d 1$.

(3) Suppose that $d \mid (\beta + 1)$. Prove that

$$n \equiv_d \sum_{i=1}^{m} (-1)^i a_i.$$

Conclude that

$$d \mid n \Leftrightarrow d \mid \sum_{i=1}^{m} a_i.$$

[XVI]

Recall that in elementary school there wasn't a "real" trick to figure out whether or not a number was divisible by 7. However, this is just a biproduct of the fact that 7 is relatively prime to 9, 10, or 11. In fact, it's the only number between 1 and 10 with this property. However, if we lived in a world where we worked base 8 instead of base 10, kids would not fear 7. For instance,

$$(3.7.2.2)_8$$

has digit sum 14 which is divisible by 7. Therefore by the exercise this number is divisible by 7. Indeed we can check this directly since (translating to base 10 where we have more confidence)

$$(3.7.2.2)_8 = 3 \cdot 8^3 + 7 \cdot 8^2 + 2 \cdot 8 + 2 = 2002 = 7 \cdot 286.$$

Exercise 3.61. Explain why the usual algorithms one learns for long division and long multiplication (in elementary school) carry over to any base.

Solution

[XIV] This corresponds to the usual (base 10) fact that either 3 or 9 divides a number iff 3 or 9 divides the sum of the digits of that number.

[XV] Recall that a two digit number is divisible by 11 iff the two digits of that number are the same. For instance,

$$11 \mid 11, 22, 33, 44, \text{ etc.}$$

However, this exercise also implies that 11 also divides $5,435,606$ since

$$(5 + 3 + 6 + 6) - (4 + 5 + 0) = 20 - 9 = 11.$$

Indeed, we can see this works since $5435606 = 494146 \cdot 11$.

Exercise 3.62. Suppose that a and b are relatively prime integers. Prove that for any integer n,

$$\gcd(ab, n) = \gcd(a, n) \cdot \gcd(b, n).$$

SOLUTION

Hint! Multiply the equations writing $\gcd(a, n)$ as an integer linear combination of a and n with an equation writing $\gcd(b, n)$ as an integer linear combination of b and n to write $\gcd(a, n) \cdot \gcd(b, n)$ as an integer linear combination of ab and n. This proves that

$$\gcd(ab, n) \mid \gcd(a, n) \cdot \gcd(b, n).$$

(without having used the fact that a and b are assumed relatively prime!) From there using the property **gcd2** you need now only prove that the product $\gcd(a, n) \cdot \gcd(b, n)$ divides ab (again, this much holds in general and is easy to prove) and n. For the harder part, you can use Exercise 3.9 applied to $a' = \gcd(a, n)$ and $b' = \gcd(b, n)$.

Exercise 3.63. Let X be a fixed set and n a fixed positive integer (we will mainly be interested in the case that $n = 2$). For any two functions

$$f : X \to \mathbb{Z}_n \quad \text{and} \quad g : X \to \mathbb{Z}_n$$

SOLUTION

define the sum and product of f and g as

$$\begin{aligned} f + g &: X \to \mathbb{Z}_n \\ &: x \mapsto f(x) + g(x) \\ f \cdot g &: X \to \mathbb{Z}_n \\ &: x \mapsto f(x) \cdot g(x). \end{aligned}$$

(1) Show that addition and multiplication of such functions is associative.

$$(f + g) + h = f + (g + h) \quad \text{and} \quad f \cdot (g \cdot h) = (f \cdot g) \cdot h.$$

(2) Now let $n = 2$ and for any such function $f : X \to \mathbb{Z}_2$ define

$$\mathcal{S}_f = \{x \in X : f(x) = [\,1\,]_2\}.$$

Conversely, if $A \subseteq X$ define a function χ_A which takes the value $[\,1\,]_2$ on points in A and $[\,0\,]_2$ on points not in A. Show that χ and $\mathcal{S}_f$ are inverse to each other.

(3) Prove that $\chi_A + \chi_B = \chi_{A \Delta B}$ where $A \,\Delta\, B$ is the symmetric difference between A and B.

(4) Prove that $\chi_A \cdot \chi_B = \chi_{A \cap B}$.

Using the above exercise we can prove that symmetric difference is associative in the sense that

$$A \,\Delta\, (B \,\Delta\, C) = (A \,\Delta\, B) \,\Delta\, C$$

as follows. By the exercise we know that {finish}

Notation 3.165. Let $\tau(n)$ denote the total number of divisors of an integer n,

Exercise 3.64. Prove that

$$\tau(n) = \prod_{p \text{ a prime}} \left(\nu_p(n) + 1\right).$$

Solution

Definition 3.166.

(1) A divisor $d \mid n$ is *complimented* if $\gcd\left(d, \frac{n}{d}\right) = 1$.

(2) An integer n is *square free* if

$$\forall m \in \mathbb{Z} \;\; m^2 \mid n \Rightarrow m = \pm 1.$$

Example 3.167.

(1) The number 6 is square free since the only integers ($\neq 1$) which divide it are ± 1, ± 2, ± 3, and ± 6– none of which are squares.

(2) The number 12 is not square free since it is divisible by $2^2 = 4$.

Exercise 3.65. Suppose that n is divisible by a distinct primes. Prove that the number of complimented divisors of n is 2^a.
Hint! Prove that the set of complimented divisors of n are

$$\left\{ \prod_{p \in P} p^{\nu_p(n)} : P \subseteq \text{Set of prime divisors of } n \right\}$$

which gives exactly one complimented divisor of n for every subset of prime divisors of n.

Solution

Exercise 3.66. For an integer n, prove that TFAE

(1) n is square free;

Solution

(2) Every divisor of n is complimented;

(3) n is not divisible by the square of any prime;

(4) Every prime divisor of n is complimented.

3.5 Matrices

3.5.1 Definitions

Definition 3.168. A ring R is commutative if we can switch the order in which elements are multiplied; that is we have that $r \cdot s = s \cdot r$ for all r and s.

With the exception of the quaternions, $\mathbb{H}$, every ring we have studied is commutative. If we are sticking with the theme of simply looking at examples, a commutative ring is one of the following rings:

$$\mathbb{Z}_n, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \text{ or } \mathbb{C}.$$

Don't forget, $\mathbb{Z} \cap \mathbb{Z}_n = \emptyset$

Convention 3.169. In this section on matrices, R is always a commutative ring.

Definition 3.170. An n by m matrix with coefficients in a R is a function $\mathbf{A}\colon \overline{n} \times \overline{m} \to R$. For a matrix $\mathbf{A}$

(1) We say that $\mathbf{A}(i, j)$ is the i, j entry of $\mathbf{A}$. Often, we will denote this with subscripts (and not bold the A) as A_{ij} instead of $\mathbf{A}(i, j)$.[I]

(2) We denote the set of all n by m matrices with coefficients in R by $\mathbf{M}_{n\times m}(R)$.

(3) If $n = m$ we say that $\mathbf{A}$ is a *square matrix* of degree n. The set of square matrices of degree n is denoted as $\mathbf{M}_n(R)$ instead of as $\mathbf{M}_{n\times n}(R)$.

This is just a fancy pants definition of a what you would expect a matrix to be, an array of elements of R with n rows and m columns. Indeed, if $\mathbf{A}$ is an n by m matrix, we write $\mathbf{A}$ as

$$\begin{pmatrix} A_{11} & A_{12} & \cdots A_{1m} \\ A_{21} & A_{22} & \cdots A_{2m} \\ \vdots & \vdots & \vdots \\ A_{n1} & A_{n2} & \cdots A_{nm} \end{pmatrix}.$$

Every entry is in R.

Exercise 3.67. Use the properties of functions to prove that two matrices are equal iff all of their coefficients are equal.

Solution

[I] This should be be familiar to you from vectors. A vector $\mathbf{v} = (v_1, v_2, v_3)$.

EXAMPLE 3.171. For instance

$$\begin{pmatrix} [2]_6 & [3]_6 & [5]_6 \\ [1]_6 & [4]_6 & [0]_6 \\ [1]_6 & [2]_6 & [1]_6 \end{pmatrix} \in \mathbf{M}_3(\mathbb{Z}_6).$$

It's $(3,2)$ entry is $[2]_6$ as highlighted:

$$\begin{pmatrix} [2]_6 & [3]_6 & [5]_6 \\ [1]_6 & [4]_6 & [0]_6 \\ [1]_6 & \boxed{[2]_6} & [1]_6 \end{pmatrix}.$$

The matrix

$$\begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

is in $\mathbf{M}_2(\mathbb{Q}), \mathbf{M}_2(\mathbb{R})$, and $\mathbf{M}_2(\mathbb{C})$ but it is not in $\mathbf{M}_2(\mathbb{Z})$ since $1/2$ is not an integer. Note that we clearly have that

$$\mathbf{M}_n(\mathbb{Z}) \subseteq \mathbf{M}_n(\mathbb{Q}) \subseteq \mathbf{M}_n(\mathbb{R}) \subseteq \mathbf{M}_n(\mathbb{C}).$$

We do not however have an inclusion such as $\mathbf{M}_2(\mathbb{Z}) \subseteq \mathbf{M}_3(\mathbb{Z})$. There are, however, several reasonable identifications of $\mathbf{M}_2(\mathbb{Z})$ with subsets of $\mathbf{M}_3(\mathbb{Z})$. For instance, it seems reasonable to identify $\mathbf{M}_2(\mathbb{Z})$ with matrices of the form

$$\begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

or one of any number of other possibilities. But, with any of these identifications we should keep in mind that the statement $\mathbf{M}_2(\mathbb{Z}) \subseteq \mathbf{M}_3(\mathbb{Z})$ is **not literally true.**

So far there's no reason to insist that addition and multiplication make sense in R. For instance, we could consider

$$\begin{pmatrix} \clubsuit & \heartsuit & \diamondsuit \\ \heartsuit & \spadesuit & \heartsuit \\ \diamondsuit & \diamondsuit & \diamondsuit \end{pmatrix}$$

as an element of $\mathbf{M}_3(\{\clubsuit, \heartsuit, \spadesuit, \diamondsuit\})$, but we will have no use for such matrices. Moreover matrices with coefficients in $\mathbb{H}$ are often studied in the literature. However, for our purposes things there are just slightly harder,[II] so we don't cover them in this introduction.

EXERCISE 3.68. Use the definition of a matrix as a function to show that two matrices are equal iff all of their entries are equal.

SOLUTION

We state the familiar definitions of addition and multiplication of matrices:

DEFINITION 3.172. Let R be a commutative ring. Then we define

(1) Let $r \in R$ and $\mathbf{A}$ be a matrix with coefficients in R. Then we define the matrix

$$(r \cdot \mathbf{A})(i,j) = r \cdot (\mathbf{A}(i,j)).$$

[II] Mostly due to the fact that determinants need to be rethought in that situation.

(2) Given two n by m matrices $\mathbf{A}$ and $\mathbf{B}$ define the sum of $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A}+\mathbf{B}$, as the matrix whose $i,\ j$ coefficient is the sum of the i and j coefficients of $\mathbf{A}$ and $\mathbf{B}$. That is

$$(\mathbf{A}+\mathbf{B})(i,j) \stackrel{\text{def}}{=} \mathbf{A}(i,j)+\mathbf{B}(i,j).$$

(3) If $\mathbf{A}$ is a n by m matrix and $\mathbf{B}$ is an m by r matrix the product of $\mathbf{A}$ and $\mathbf{B}$, denoted $\mathbf{A}\cdot\mathbf{B}$, is the n by r matrix whose components are

$$(\mathbf{A}\cdot\mathbf{B})(i,j) \stackrel{\text{def}}{=} \sum_{k=1}^{m} \mathbf{A}(i,k)\cdot\mathbf{B}(k,j)$$

where the products on the RHS are defined in R.

If we think of an n by m matrix as an element of $(R^n)^m$, addition and scalar multiplication correspond to the usual component-wise versions. However, the multiplication is defined in a more complicated fashion. It will turn out that it is defined in such a way so that the multiplication of matrices corresponds to the composition of linear functions.

REMARK 3.173. One way to think of matrix multiplication, if you haven't seen it before, is to imagine it as a sequence of dot products. In more detail, if we have a matrix $\mathbf{A}$ with row m-vectors and matrix $\mathbf{B}$ with column m-vectors given by

$$\mathbf{A}=\begin{pmatrix}\mathbf{r}(\mathbf{A})_1\\ \vdots\\ \mathbf{r}(\mathbf{A})_n\end{pmatrix} \qquad \mathbf{B}=\begin{pmatrix}\mathbf{c}(\mathbf{B})_1 \mid \cdots \mid \mathbf{c}(\mathbf{B})_r\end{pmatrix}$$

then their product $\mathbf{A}\cdot\mathbf{B}$ has $i,\ j$ entry given by the dot product $\mathbf{r}_i \bullet \mathbf{c}_j$, i.e.

$$(\mathbf{A}\cdot\mathbf{B})(i,j)=\mathbf{r}(\mathbf{A})_i \bullet \mathbf{c}(\mathbf{B})_j$$

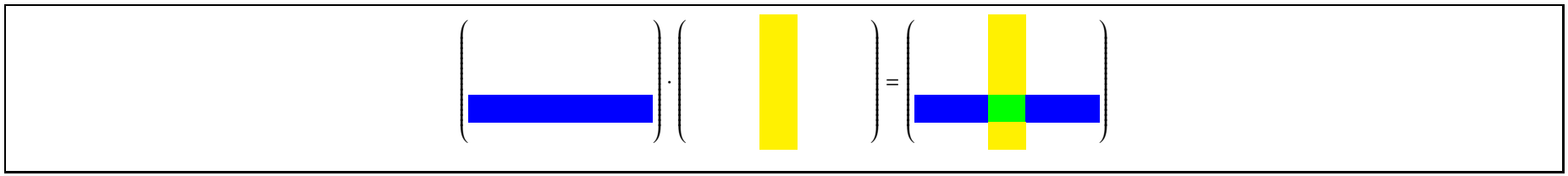


Figure 3.4: YELLOW and BLUE make GREEN. The entry of the product is the dot product of the blue row with the yellow column and is positioned at the green location.

EXAMPLE 3.174.

(1) I hope Figure 3.4 helps you keep track of how to do matrix multiplication. In practice:

$$\begin{pmatrix} \mathbf{3} & \mathbf{1} & \mathbf{1} \\ 2 & 1 & 6 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{1} & 5 \\ \mathbf{7} & 3 \\ \mathbf{1} & 1 \end{pmatrix} (1,1) = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 7 \\ 1 \end{bmatrix} = 3 + 7 + 1 = 11 \qquad \begin{pmatrix} 11 & \\ & \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{3} & \mathbf{1} & \mathbf{1} \\ 2 & 1 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & \mathbf{5} \\ 7 & \mathbf{3} \\ 1 & \mathbf{1} \end{pmatrix} (1,2) = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \bullet \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix} = 15 + 3 + 1 = 19 \qquad \begin{pmatrix} 11 & 19 \\ & \end{pmatrix}$$

$$\begin{pmatrix} 3 & 1 & 1 \\ \mathbf{2} & \mathbf{1} & \mathbf{6} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{1} & 5 \\ \mathbf{7} & 3 \\ \mathbf{1} & 1 \end{pmatrix} (2,1) = \begin{bmatrix} 2 \\ 1 \\ 6 \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 7 \\ 1 \end{bmatrix} = 2 + 7 + 6 = 15 \qquad \begin{pmatrix} 11 & 19 \\ 15 & \end{pmatrix}$$

$$\begin{pmatrix} 3 & 1 & 1 \\ \mathbf{2} & \mathbf{1} & \mathbf{6} \end{pmatrix} \cdot \begin{pmatrix} 1 & \mathbf{5} \\ 7 & \mathbf{3} \\ 1 & \mathbf{1} \end{pmatrix} (1,2) = \begin{bmatrix} 2 \\ 1 \\ 6 \end{bmatrix} \bullet \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix} = 10 + 3 + 6 = 19 \qquad \begin{pmatrix} 11 & 19 \\ 15 & 19 \end{pmatrix}.$$

(2) For most students, the only thing that may may be new here is how to multiply matrices that aren't numbers. In practice this shouldn't be any burden, simply follow the same rules as you may be familiar with from multiplying matrices, but remember to perform the multiplication in the ring R. For instance:

$$\begin{pmatrix} [4]_6 & [2]_6 \\ [3]_6 & [5]_6 \end{pmatrix} \cdot \begin{pmatrix} [1]_6 & [3]_6 \\ [2]_6 & [4]_6 \end{pmatrix} = \begin{pmatrix} [4\cdot 1 + 2\cdot 2]_6 & [4\cdot 3 + 2\cdot 4]_6 \\ [3\cdot 1 + 5\cdot 2]_6 & [3\cdot 3 + 5\cdot 4]_6 \end{pmatrix} = \begin{pmatrix} [8]_6 & [20]_6 \\ [13]_6 & [29]_6 \end{pmatrix} = \begin{pmatrix} [2]_6 & [2]_6 \\ [1]_6 & [5]_6 \end{pmatrix}.$$

(3) For any ring besides the trivial ring $\mathbb{Z}_1$, matrix multiplication of even two by two matrices is never (in general) commutative. For instance the matrices

$$\mathfrak{L} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \ \mathfrak{R} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

satisfy

$$\mathfrak{L} \cdot \mathfrak{R} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \mathfrak{R} \cdot \mathfrak{L} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

which are not equal since for any ring (besides $\mathbb{Z}_1$) $2 \neq 1$.[III]

Note that any two matrices in $\mathbf{M}_n(R)$ can both be added or multiplied (hence they fit into the theme of the chapter)!

DEFINITION 3.175. **(1)** For any positive integer n, $\mathbf{O} = \mathbf{O}_n$ denotes the n by n zero matrix, i.e. the matrix all of whose coefficients are 0.

(2) For any positive integer n, $\mathbf{I} = \mathbf{I}_n$ is the n by n identity matrix, i.e., the matrix with coefficients $\mathbf{I}_{i,j} = 1$ if $i = j$ and equal to 0 otherwise. That is $\mathbf{I}(i, j) = 1 \Leftrightarrow i = j$.

Technically the matrices $\mathbf{I}_n$ and $\mathbf{O}_n$ depend on the ring as well as the integer n.

[III] If it did, then

$$2 = 1 \Rightarrow 1 + 1 = 1 \Rightarrow 1 = 0 \Rightarrow \forall r \in R, r = r \cdot 1 = r \cdot 0 = 0.$$

Example 3.176. If $R = \mathbb{Z}$, then

$$\mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

If $R = \mathbb{Z}_n$, then

$$\mathbf{I}_3 = \begin{pmatrix} [\,1\,]_n & [\,0\,]_n & [\,0\,]_n \\ [\,0\,]_n & [\,1\,]_n & [\,0\,]_n \\ [\,0\,]_n & [\,0\,]_n & [\,1\,]_n \end{pmatrix}.$$

The following theorem says that, in many respects, matrices behave like the rings we're already familiar with including

$$\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_n, \text{ and } \mathbb{H}$$

with the identity matrix playing the role of the number 1 and the zero matrix playing the role of the number 0. We state it only for square matrices. Tweaking the statements to the non-square case is left to you.

Theorem 3.177. *For square matrices* $\mathbf{A}$, $\mathbf{B}$, *and* $\mathbf{C}$ *of the same degree,* n, *we have*

(1) (Identity Properties)

$$\mathbf{I}_n \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{I}_n = \mathbf{A}.$$
$$\mathbf{O}_n + \mathbf{A} = \mathbf{A} + \mathbf{O}_n = \mathbf{A}.$$

(2) (Inverse Properties) *There exists a matrix* $-\mathbf{A}$ *so that*

$$\mathbf{A} + (-\mathbf{A}) = (-\mathbf{A}) + \mathbf{A} = \mathbf{O}_n.$$

(3) (Distributive Properties)

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$$
$$(\mathbf{A} + \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot \mathbf{C} + \mathbf{B} \cdot \mathbf{C}$$

(4) (Associative Properties)

$$(\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C})$$
$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}).$$

(5) (Commutative Properties)

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}.$$

With the exception of the commutivity of multiplication (or lack thereof in our present situation), these are the same properties that held in Theorem 3.51.

Proof. We show only two part of this.

The matrix $\mathbf{A} \cdot \mathbf{I}_n$ has i,k coefficient equal to

$$\begin{aligned}(\mathbf{A} \cdot \mathbf{I}_n)(i,k) &= \sum_{k=1}^{n} \mathbf{A}(i,j) \cdot \mathbf{I}(j,k) \\ &= \mathbf{A}(i,1) \cdot \cancel{\mathbf{I}(1,k)} + \cdots + \mathbf{A}(i,k-1) \cdot \cancel{\mathbf{I}(k-1,k)} + \mathbf{A}(i,k) \cdot \cancel{\mathbf{I}(k,k)}^{1} + \mathbf{A}(i,k+1) \cdot \cancel{\mathbf{I}(k+1,k)} + \cdots + \mathbf{A}(i,n) \cdot \cancel{\mathbf{I}(n,k)} \\ &= 0 + \cdots + 0 + \mathbf{A}(i,k) + 0 + \cdots + 0 \\ &= \mathbf{A}(i,k).\end{aligned}$$

That is the i,k coefficient of the product is equal to the i,k coefficient of $\mathbf{A}$. Hence, every coefficient of this product and the matrix $\mathbf{A}$ coincide and the two are equal. To prove that $\mathbf{I}_n \cdot \mathbf{A} = \mathbf{A}$ is the same.

The commutativity of addition, the associativity of addition are obvious and follow from the corresponding properties of addition in R^{n^2}. The associativity of multiplication can be proven now without too much effort, but will be proven later by turning matrices into (linear) functions.

$$\begin{aligned}(\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}))(i,j) &= \sum_{k=1}^{m} \mathbf{A}(i,k) \cdot (\mathbf{B} + \mathbf{C})(k,j) \\ &= \sum_{k=1}^{m} \mathbf{A}(i,k) \cdot (\mathbf{B}(k,j) + \mathbf{C}(k,j)) \\ &= \sum_{k=1}^{m} (\mathbf{A}(i,k) \cdot \mathbf{B}(k,j) + \mathbf{A}(i,k) \cdot \mathbf{C}(k,j)) \\ &= \sum_{k=1}^{m} \mathbf{A}(i,k) \cdot \mathbf{B}(k,j) + \sum_{k=1}^{m} \mathbf{A}(i,k) \cdot \mathbf{C}(k,j) \\ &= (\mathbf{A} \cdot \mathbf{B})(i,j) + (\mathbf{A} \cdot \mathbf{C})(i,j)\end{aligned}$$

as desired. The proof of the other distributive law is the same. ■

As pointed out in the proof, we could use the same method to prove that other properties of matrix multiplication and addition hold, including the associative property of multiplication (which for our purposes that is the most important).[IV] However, we will see a more elegant proof of this which turns matrices into functions.

3.5.2 Linear Transformations*

Recall that R^n is the set of ordered n–tuples of elements of R which we will typically write as a column vector. For instance an element of $\mathbb{Z}_5^3$ would be $\begin{bmatrix}[3]_5 \\ [2]_5 \\ [1]_5\end{bmatrix}$. Recall the

[IV] In fact, for any ring R, $\mathbf{M}_n(R)$ forms a ring which is non-commutative for $n \geq 2$. We won't use this in part of the main text.

definitions of scalar multiplication, and addition of elements of R^n given in Appendix A. For example, if $c = [\,3\,]_5$ and $\mathbf{v} = \begin{bmatrix}[\,3\,]_5\\ [\,2\,]_5\\ [\,1\,]_5\end{bmatrix}$, as above, we would have

$$\begin{aligned} c \cdot \mathbf{v} &= [\,3\,]_5 \cdot \begin{bmatrix}[\,3\,]_5\\ [\,2\,]_5\\ [\,1\,]_5\end{bmatrix} \\ &= \begin{bmatrix}[\,3\,]_5 \cdot [\,3\,]_5\\ [\,3\,]_5 \cdot [\,2\,]_5\\ [\,3\,]_5 \cdot [\,1\,]_5\end{bmatrix} \\ &= \begin{bmatrix}[\,4\,]_5\\ [\,1\,]_5\\ [\,3\,]_5\end{bmatrix}. \end{aligned}$$

If in addition $\mathbf{w} = \begin{bmatrix}[\,1\,]_5\\ [\,0\,]_5\\ [\,4\,]_5\end{bmatrix}$, then

$$\begin{aligned} \mathbf{v} + \mathbf{w} &= \begin{bmatrix}[\,3\,]_5\\ [\,2\,]_5\\ [\,1\,]_5\end{bmatrix} + \begin{bmatrix}[\,1\,]_5\\ [\,0\,]_5\\ [\,4\,]_5\end{bmatrix} \\ &= \begin{bmatrix}[\,3\,]_5 + [\,1\,]_5\\ [\,2\,]_5 + [\,0\,]_5\\ [\,1\,]_5 + [\,4\,]_5\end{bmatrix} \\ &= \begin{bmatrix}[\,4\,]_5\\ [\,2\,]_5\\ [\,0\,]_5\end{bmatrix}. \end{aligned}$$

Definition 3.178. A function

$$F : R^m \to R^n$$

is said to be a *linear mapping* (or a *linear function* or a *linear transformation*) if

$$\forall \mathbf{v}, \mathbf{w} \in R^n, \qquad F(\mathbf{v} + \mathbf{w}) = F(\mathbf{v}) + F(\mathbf{w}) \tag{L1}$$

$$\forall \mathbf{v} \in R^n, c \in R, \qquad F(c\,\mathbf{v}) = c\,F(\mathbf{v}). \tag{L2}$$

When **L1** holds for a function F, we say that F *preserves addition*, and when **L2** holds for a function F, we say that F *preserves scalar multiplication.* Therefore, a function is linear whenever it preserves both addition and scalar multiplication.

Example /Theorem 3.179. The composition of two linear transformations (when defined) is a linear transformation.

Proof. Suppose that F and G are linear transformations so that $F \circ G$ makes sense (i.e. so that the second set of G is the first set of F). Let $\mathbf{v}$ and $\mathbf{w}$ be arbitrary elements of the first set (the domain) of G. Let $\mathbf{u}'$ and $\mathbf{v}'$ denote $G(\mathbf{u})$ and $G(\mathbf{v})$ respectively. Then

$$\begin{aligned}
(F \circ G)(\mathbf{v} + \mathbf{w}) &= F(G(\mathbf{u} + \mathbf{v})) && \text{(Def'n of comp of functions.)} \\
&= F(G(\mathbf{u}) + G(\mathbf{v})) && \text{(Since L1 holds for } G\text{.)} \\
&= F(\mathbf{u}' + \mathbf{v}') && \text{(Def'n of } \mathbf{u}' \text{ and } \mathbf{v}'\text{.)} \\
&= F(\mathbf{u}') + F(\mathbf{v}') && \text{(Since L1 holds for } F\text{.)} \\
&= F(G(\mathbf{u})) + F(G(\mathbf{v})) && \text{(Def'n of } \mathbf{u}' \text{ and } \mathbf{v}'\text{.)} \\
&= (F \circ G)(\mathbf{u}) + (F \circ G)(\mathbf{v}) && \text{(Def'n of comp of functions.)}
\end{aligned}$$

Therefore we have $F \circ G$ satisfies L1 as well, i.e. $F \circ G$ preserves addition.

Now we must show that $F \circ G$ preserves scalar multiplication assuming that both F and G do. To that end let $\mathbf{v}$ be an element of the first set (the domain) of G and $c \in R$ be arbitrary. Then since G satisfies L2 we have that $G(c\mathbf{v}) = cG(\mathbf{v})$. Now since F also satisfies L2 we have that

$$F(G(c \cdot \mathbf{v}) = F(c \cdot G(\mathbf{v})) = c \cdot F(G(\mathbf{v}).$$

But this means that $(F \circ G)(c \cdot \mathbf{v}) = c \cdot (F \circ G)(\mathbf{v})$ and as such the composition also satisfies L2, i.e. preserves scalar multiplication. ■

Remark 3.180. Since the focus of mathematics is as much (probably more so ... a lot more so) why something is true as opposed to just what is true I should stress that when proving that $F \circ G$ preserves addition we only used the assumption that both F and G preserve addition. Similarly, in proving that $F \circ G$ preserves scalar multiplication we only used the assumptions that F and G do. Therefore we have really proven that the composition of two functions which preserve addition also preserves additions, and the composition of two functions which preserve scalar multiplication also preserves scalar multiplication.

Exercise 3.69. Let F and G be linear functions.

(1) Show that $F + G$ is also a linear function. (Naturally $F + G$ is defined to be the function which takes a vector $\mathbf{v}$ to $F(\mathbf{v}) + G(\mathbf{v})$.)

(2) If in addition we assume that F is a bijection, show that F^{-1} is also linear.

SOLUTION

If we think of R^n as n by 1 matrices (i.e. as column vectors) for any n by n matrix $\mathbf{A}$, $\mathbf{A} \cdot v$ is defined and is equal to another n by 1 matrix, i.e. an element of this same starting set R^n. Therefore every matrix $\mathbf{A}$ determines a linear transformation $\mathbf{A}^\flat$. For instance if

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then

$$\mathbf{A}^\flat(x, y) = (ax + by, cx + dy)$$

since

$$\mathbf{A} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$$

In practice figuring out the function is always this easy. Simply read across the rows and put $+$ signs and variables where appropiate. Let's be a bit more precise. If $\mathbf{A}$ is an n by m matrix (n rows and m columns!) with entries $A_{i,j}$, then we are defining

$$\mathbf{A}(x_1, \dots x_m) = (u_1, \dots, u_n) \quad \text{where} \quad u_i = \sum_{j=1}^{m} A_{i,j} x_j.$$

Theorem 3.181. *Under this correspondence, the multiplication of matrices corresponds to the composition of functions. More precisely if* $\mathbf{A}$ *and* $\mathbf{B}$ *a matrices which can be multiplied, then*

$$(\mathbf{A} \cdot \mathbf{B})^\flat = \mathbf{A}^\flat \circ \mathbf{B}^\flat.$$

Proof. This isn't hard it simply takes a bit of focus. Just play with coefficients {Do I want to make this better? }. ■

How do we go backwards from a linear transformation to a matrix? For example, if $F\colon R^2 \to R^2$ is a linear transformation, then

$$F(x,y) = F(x\mathbf{e}_1 + y\mathbf{e}_2) = F(x\mathbf{e}_1) + F(y\mathbf{e}_2) = x \cdot F(\mathbf{e}_1) + y \cdot F(\mathbf{e}_2).$$

This shows that as soon as we figure out what F does to the standard basis vectors $\mathbf{e}_1 = (1,0)$ and $\mathbf{e}_2 = (0,1)$ we know what F does to every point (x,y). This clearly generalizes

Lemma 3.182. *Any linear transformation* $F\colon R^n \to R^m$ *is determined by what it does to the standard basis*

$$\mathbf{e}_1, \dots, \mathbf{e}_n$$

if $G\colon R^n \to R^m$ *is another linear transformation that is equal to* F *when restricted to the standard basis, then* $F = G$.

Remark 3.183. This is obviously nonsense for non-linear functions. For instance the functions $f(x) = x$ and $f(x) = x^2$ both do the same thing to the only standard basis vector $1 \in \mathbb{R}$, but they are clearly not the same function.

Definition 3.184. Let $F\colon R^n \to R^m$ be a linear transformation. Define $F^\sharp$ as the matrix whose i^{th} column vector is $F(\mathbf{e}_i)$.

In more detail, $F^\sharp$ is the matrix $\mathbf{A}$ with coefficients A_{ij} iff

$$\begin{aligned} F(x_1, \dots, x_n) &= x_1 F(\mathbf{e}_1) + \cdots x_n F(\mathbf{e}_n) \\ &= x_1 (A_{11}\mathbf{e}_1 + \cdots A_{m1}\mathbf{e}_m) + \cdots x_n (A_{1n}\mathbf{e}_1 + \cdots A_{mn}\mathbf{e}_m) \\ &= (x_1 A_{11} + \cdots x_n A_{1n})\,\mathbf{e}_1 + \cdots (x_1 A_{m1} + \cdots x_n A_{mn})\,\mathbf{e}_m \\ &= (x_1 A_{11} + \cdots x_n A_{1n}, \dots, x_1 A_{m1} + \cdots x_n A_{mn})\,. \end{aligned}$$

Example 3.185. Fix a complex number $z = a + bi$. Complex multiplication by z, m_z, determines a linear map (with $\mathbf{e}_1 = 1$ and $\mathbf{e}_2 = i$). We have that

$$m_z(\mathbf{e}_1) = z \cdot \mathbf{e}_1 = a + bi = a\mathbf{e}_1 + b\mathbf{e}_2 \quad \text{and} \quad m_z(\mathbf{e}_2) = z \cdot \mathbf{e}_2 = ai - b = -b\mathbf{e}_1 + a\mathbf{e}_2.$$

Therefore

$$m_z{}^\sharp = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Example 3.186. The derivative is a linear function from on the set of differentiable functions. If we restict this to simply look at the set of polynomials of degree less than 5, then we can identify $\mathbf{e}_1$ with 1, $\mathbf{e}_2$ with x, $\mathbf{e}_3$ with x^2, $\mathbf{e}_4$ with x^3, $\mathbf{e}_5$ with x^4 and $\mathbf{e}_6$ with x^5. Thus, we are identifying a polynomial of degree five or less with $\mathbb{R}^6$. With this identification, we have that the derivative does the following to these standard basis vectors:

$$\mathbf{e}_1 \mapsto 0 \qquad \mathbf{e}_2 \mapsto \mathbf{e}_1$$
$$\mathbf{e}_3 \mapsto 2\mathbf{e}_2 \qquad \mathbf{e}_4 \mapsto 3\mathbf{e}_3$$
$$\mathbf{e}_5 \mapsto 4\mathbf{e}_4 \qquad \mathbf{e}_6 \mapsto 5\mathbf{e}_5.$$

Thus, the 6 by 6 matrix corresponding to this linear transformation is

$$\mathbf{Der} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Theorem 3.187. *The operations $\flat$ and $\sharp$ are inverse to each other. That is for any linear transformation F we have that $\left(F^{\sharp}\right)^{\flat} = F$ and for any matrix $\mathbf{A}$ we have that $\left(\mathbf{A}^{\flat}\right)^{\sharp} = \mathbf{A}$.*

Proof. For any matrix $\mathbf{A}$ we have that $\mathbf{A} \cdot \mathbf{e}_i$ is the i^{th} column of $\mathbf{A}$. Therefore $\left(\mathbf{A}^{\flat}\right)^{\sharp} = \mathbf{A}$. Conversely, for any linear transformation F we have that $F^{\sharp\flat}$ is the linear transformation which takes $\mathbf{e}_i$ to $F(\mathbf{e}_i)$. But by the lemma, Lemma 3.182 says that there is only one such linear transformation, namely F itself. Therefore $\left(F^{\sharp}\right)^{\flat} = F$ as desired. ∎

Corollary 3.188. *For any two composable linear transformations F and G we have that*

$$(F \circ G)^{\sharp} = F^{\sharp} \cdot G^{\sharp}.$$

Proof. By the theorem, Theorem 3.187, we have that $\left(F^{\sharp}\right)^{\flat} = F$ and $\left(G^{\sharp}\right)^{\flat} = G$. Thus

$$\begin{aligned} (F \circ G)^{\sharp} &= \left((F^{\sharp})^{\flat} \circ (G^{\sharp})^{\flat}\right)^{\sharp} \\ &= \left(\left(F^{\sharp} \cdot G^{\sharp}\right)^{\flat}\right)^{\sharp} && \text{(Theorem 3.181)} \\ &= F^{\sharp} \cdot G^{\sharp}. && \text{(The theorem, Theorem 3.187)} \end{aligned}$$

∎

You should note that we proved this property about $\sharp$, not by going back to the definition, but by using the properties that we have already proved about $\flat$ and $\sharp$ (namely that they are inverses to each other and that $\flat$ has the corresponding property.) Similarly, we will prove the following corollary not by going back to the definition, but by using the link that we have no established between linear functions and composition and matrices and matrix multiplication.

Corollary 3.189. *Matrix multiplication is associative.*

Proof. Let $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ be square matrices of the same size. Then, since $\mathbf{A}^\flat$, $\mathbf{B}^\flat$, and $\mathbf{C}^\flat$ are linear functions, and hence function, their product is associative. Then we have

$$\begin{aligned}(\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C}))^\flat &= \mathbf{A}^\flat\circ(\mathbf{B}\cdot\mathbf{C})^\flat\\ &= \mathbf{A}^\flat\circ(\mathbf{B}^\flat\circ\mathbf{C}^\flat)\\ &= (\mathbf{A}^\flat\circ\mathbf{B}^\flat)\circ\mathbf{C}^\flat\\ &= (\mathbf{A}\cdot\mathbf{B})^\flat\circ\mathbf{C}^\flat\\ &= ((\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C})^\flat\,.\end{aligned}$$

Taking this equality and applying $\sharp$ shows us that

$$\begin{aligned}(\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C}))^\flat = ((\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C})^\flat &\Rightarrow \left(\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})^\flat\right)^\sharp = \left(((\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C})^\flat\right)^\sharp\\ &\Rightarrow \left(((\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C})^{\cancel{\flat}}\right)^{\cancel{\sharp}} = \left(\left(\mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C})^{\cancel{\flat}}\right)\right)^{\cancel{\sharp}} \quad (\flat \text{ and } \sharp \text{ cancel})\\ &\Rightarrow \mathbf{A}\cdot(\mathbf{B}\cdot\mathbf{C}) = (\mathbf{A}\cdot\mathbf{B})\cdot\mathbf{C}\end{aligned}$$

as desired. ∎

Definition 3.190. A square matrix $\mathbf{A}$ is invertible if there exists a matrix $\widehat{\mathbf{A}}$ so that

$$\mathbf{A}\cdot\widehat{\mathbf{A}} = \widehat{\mathbf{A}}\cdot\mathbf{A} = \mathbf{I}.$$

Remark 3.191. If $\widehat{\mathbf{A}}'$ and $\widehat{\mathbf{A}}$ are two inverses for the matrix $\mathbf{A}$, then

$$\widehat{\mathbf{A}} = \widehat{\mathbf{A}}\cdot\left(\mathbf{A}\cdot\widehat{\mathbf{A}}'\right) = \left(\widehat{\mathbf{A}}\cdot\mathbf{A}\right)\cdot\widehat{\mathbf{A}}' = \widehat{\mathbf{A}}'.$$

Hence if a matrix has one inverse, it has only one inverse. Using essentially the same proof, we will see (using only the associative property) in Theorem 4.34 that this holds in a much broader context. Therefore, in this situation it makes sense to define $\mathbf{A}^{-1}$ as the $\widehat{\mathbf{A}}$ in the definition. We are also lucky in the case of square matrices that we only need to check one of the two inequalities since if we had $\mathbf{A}\cdot\widehat{\mathbf{A}} = \mathbf{I}$, then by a theorem about matrices, we automatically have $\widehat{\mathbf{A}}\cdot\mathbf{A} = \mathbf{I}$.

Exercise 3.70.

SOLUTION

(1) Show that $(\mathbf{I}_n)^\flat = \mathbf{1}_{R^n}$ and $(\mathbf{1}_{R^n})^\sharp = \mathbf{I}_n$ (since $\flat$ and $\sharp$ are inverse bijections, you only need to show one).

(2) Show that a matrix $\mathbf{A}$ is invertible iff the linear transformation $\mathbf{A}^\flat$ is an invertible function. Conversely, show that F is an invertible linear function iff $F^\sharp$ is an invertible matrix.

Hint! Apply $\flat$ to the equality $\mathbf{A}\cdot\mathbf{B} = \mathbf{I}_n$ and use the previous part.

The following definitions rank among the top 5 definitions in all of linear algebra. In this class they will play a reduced role. In fact we only mention them in a hand full of exercises in this chapter on linear algebra.

Definition 3.192. Let $\mathcal{V}$ denote a set $\{\mathbf{v}_1, \dots, \mathbf{v}_m\} \subseteq R^n$ be a set of m, n-vectors.

(1) The set $\mathcal{V}$ is said to be *linearly dependent* if there exists scalars $r_1, \dots, r_m \in R$, not all zero, so that

$$\mathbf{0} = r_1\mathbf{v}_1 + \cdots r_m\mathbf{v}_m.$$

(2) The span of $\mathcal{V}$, denoted $\langle\, \mathcal{V} \,\rangle$, is the set of linear combinations of the elements of $\mathcal{V}$. That is it is the set

$$\{r_1\mathbf{v}_1 + \cdots r_n\mathbf{v}_n : r_1, \dots, r_n \in R\}.$$

If $\langle\, \mathcal{V} \,\rangle = R^n$, then we say that $\mathcal{V}$ is a *spanning set*.

(3) The set $\mathcal{V}$ is called a *basis* if it is linearly independent and a spanning set.

Exercise 3.71.

(1) Prove that every subset of a linearly independent set is linearly independent.

(2) Prove that every subset of a non spanning set is non spanning.

(3) Prove that every superset of a linearly dependent set is linearly dependent.

(4) Prove that every superset of a non spanning set is non spanning.

Solution

Exercise 3.72. Suppose that $\mathbb{F}$ is a field. For $\mathcal{V} := \{\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_m\} \subseteq \mathbb{F}^n$, prove that

(1) Prove that $\mathcal{V}$ is linearly dependent iff there exists a $i \in \overline{n}$ so that

$$\mathbf{v}_i \in \langle\, \mathcal{V} \,\rangle$$

(2) Prove that

$$\mathbf{v} \in \langle\, \mathcal{V} \,\rangle \Rightarrow \langle\, \mathcal{V} \,\rangle = \langle\, \mathcal{V} \cup \mathbf{v} \,\rangle\,.$$

(3) Prove that every spanning set contains a basis.

(4) Prove that $\mathcal{V}$ is a basis iff every $\mathbf{v} \in \mathbb{F}^n$ can be written uniquely as a linear combination of elements of $\mathcal{V}$.

(5) Prove that $\mathcal{V}$ is linearly independent iff the matrix

$$(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_m)$$

is injective.

Solution

(6) Prove that $\operatorname{Im}(() \mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n) = \langle ( ) \mathcal{V} \rangle$.

(7) Prove that $\mathcal{V}$ is a spanning set iff the matrix

$$(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n)$$

is surjective.

(8) Prove that $\mathcal{V}$ is a basis iff the matrix

$$(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n)$$

is bijective.

Definition 3.193. A set $V \subseteq R^n$ is said to be a *sub-space* if

(1) For all $\mathbf{v}, \mathbf{w} \in V$, $\mathbf{v} + \mathbf{w} \in V$.

(2) For all $(1) \in R$ and $\mathbf{v} \in V$, $c \cdot \mathbf{v} \in V$.

Exercise 3.73. Suppose $\mathrm{F}: R^m \to R^n$ is a linear transformation.

Solution

(1) Prove that

$$\operatorname{Ker}(\mathrm{F}) \stackrel{\text{def}}{=} \{\mathbf{v} \in R^n : \mathrm{F}(\mathbf{v}) = \mathbf{0}\}$$

is a subspace of V.

(2) Prove that $\operatorname{Im}(\mathrm{F})$ (defined as always) is a subspace of R^m.

(3) For any subset $\mathcal{V} \subseteq R^n$ prove that $\langle \mathcal{V} \rangle$ is a subspace.

3.5.3 The Determinant

There are a few ways one can think of the determinant.

3.5.3.1 Method One: Inductive

Given an n by n matrix $\mathbf{A}$ and $(k, l) \in \overline{n} \times \overline{m}$ define $\mathbf{A}^{(k,l)}$ as the $n-1$ by $n-1$ matrix

$$\mathbf{A}^{(k,l)}(i,j) = \begin{cases} \mathbf{A}(i,j) & \text{if } i < k \text{ and } j < l \\ \mathbf{A}(i+1,j) & \text{if } i \geq k \text{ and } j < l \\ \mathbf{A}(i,j+1) & \text{if } i < k \text{ and } j \geq l \\ \mathbf{A}(i+1,j+1) & \text{if } i \geq k \text{ and } j \geq l. \end{cases}$$

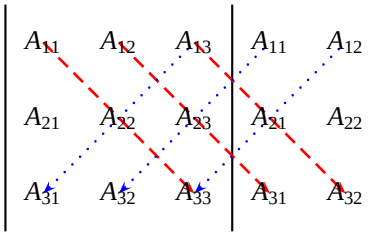
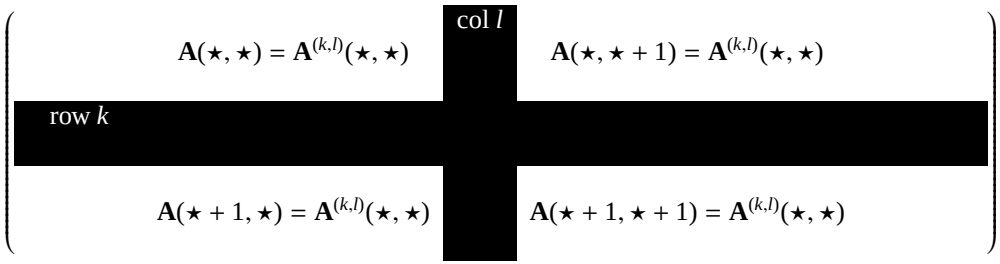


Figure 3.5: The figure gives the standard "trick" to remember how to take the determinant of a 3 by 3 matrix. Read product of terms along lines. Red lines (lines going forward) get a plus sign and blue lines (lines going backwards) getting a negative sign. In this way the determinant of this matrix is

$$A_{11}A_{22}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{21}A_{32} - A_{13}A_{22}A_{31} - A_{11}A_{23}A_{32} - A_{12}A_{21}A_{33}.$$

This is just a fancy way of saying that $\mathbf{A}^{(k,l)}$ is the matrix one would get if we covered up the k^{th} row and l^{th} column.



EXAMPLE 3.194. Let $\mathbf{A}$ denote the 4 by 4 (integer) matrix

$$\begin{pmatrix} 3 & 2 & 1 & 4 \\ 1 & 2 & 2 & 1 \\ 3 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

then $\mathbf{A}^{(2,3)}$ is the 3 by 3 matrix

$$\begin{pmatrix} 3 & 2 & \blacksquare & 4 \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ 3 & 1 & \blacksquare & 1 \\ 1 & 2 & \blacksquare & 4 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 4 \\ 3 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}$$

and $\mathbf{A}^{(4,2)}$ is the 3 by 3 matrix

$$\begin{pmatrix} 3 & \blacksquare & 1 & 4 \\ 1 & \blacksquare & 2 & 1 \\ 3 & \blacksquare & 1 & 1 \\ \blacksquare & \blacksquare & \blacksquare & \blacksquare \end{pmatrix} = \begin{pmatrix} 3 & 1 & 4 \\ 1 & 2 & 1 \\ 3 & 1 & 1 \end{pmatrix}$$

Definition 3.195. (**Inductive**) We define the *determinant* of an n by n matrix inductively as follows:

Def'n 1 of 3

(1) The determinant of a 1 by 1 matrix (r) equals r.

(2) Inductively, assume we can find the determinant of a $n-1$ by $n-1$ matrix. Then, we define the determinant of an n by n matrix $\mathbf{A}$ as

$$\det \mathbf{A} = \sum_{k=1}^{n} (-1)^{1+k} \mathbf{A}(1,k) \cdot \det\left(\mathbf{A}^{(1,k)}\right).$$

(This is known as the cofactor expansion of the determinant along the first row.)

Notice that this definition makes perfectly good sense because each of the $\det\left(\mathbf{A}^{(1,k)}\right)$ terms which show up denote the determinant of an $n-1$ by $n-1$ matrix, which by assumption, has been defined already.

Remark 3.196. We have defined the determinant in terms of expansion about the first row. We could have also defined the determinant by expanding about any row

$$\det \mathbf{A} = \sum_{k=1}^{n} (-1)^{i+k} \mathbf{A}(i,k) \cdot \det\left(\mathbf{A}^{(i,k)}\right)$$

or any column

$$\det \mathbf{A} = \sum_{i=1}^{n} (-1)^{i+k} \mathbf{A}(i,k) \cdot \det\left(\mathbf{A}^{(i,k)}\right).$$

These facts follow easily from some general theory, but we omit this theory and their proofs. They can however be verified directly by using the formulas for the special cases that $n = 2$ or $n = 3$. Roughly, the proof relies on the subtle but obvious fact that

$$\left(\mathbf{A}^{(i,k)}\right)^{(i,l)} = \left(\mathbf{A}^{(i,l+1)}\right)^{(i,j)}$$

whenever $j < l$.

The two cases we care most about are the $n = 2$ and the $n = 3$ cases. There we see that for a 2 by 2 matrix

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

we have that

$$\begin{aligned} \det \mathbf{A} &= A_{11} \det((A_{22})) - A_{12} \det((A_{21})) \\ &= A_{11} A_{22} - A_{12} A_{21} \end{aligned}$$

where we have made use of the part the definition which says that the determinant of a 1 by 1 matrix is simply the entry of that matrix. For a 3 by 3 matrix

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

the determinant of $\mathbf{A}$ is given by

$$\det \mathbf{A} = A_{11} \det \begin{pmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{pmatrix} - A_{12} \det \begin{pmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{pmatrix} + A_{13} \det \begin{pmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix}$$

and using the formula just derived for the determinant of a 2 by 2 matrix:

$$\begin{aligned} &= A_{11}\left(A_{22}A_{33} - A_{23}A_{33}\right) - A_{12}\left(A_{21}A_{33} - A_{23}A_{31}\right) + A_{13}\left(A_{21}A_{32} - A_{22}A_{31}\right) \\ &= A_{11}A_{22}A_{33} - A_{11}A_{23}A_{33} - A_{12}A_{21}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{21}A_{32} - A_{13}A_{22}A_{31}. \end{aligned}$$

Now, at least in theory, we could use our formula to calculate the determinant of a 4 by 4 matrix. Note however, this formula will have $24 = 4 \cdot 6$ (4 columns and 6 entries for each 3 by 3 determinant) terms in it. Needless to say, this gets ugly rather quickly. Note also, in order to calculate the determinant of, say, a 7 by 7 matrix, we'd need the general formula for how to calculate the determinant of 7, 6 by 6 determinants, which could take a while!

Exercise 3.74. Using this method, prove that for every positive integer n, the determinant of the n by n identity matrix, $\mathbf{I}_n$, is 1.

Solution

3.5.3.2 Method Two: Axiomatic

Interestingly enough, the definition which leads to the fastest calculations and is often the best way of proving things is this one. However, it is an axiomatic definition (see introduction to the notes), and as such, students initially tend to avoid it like the plague.

Throughout this section we fix n to be some natural number, and R some fixed commutative ring. To make the notation appear a bit cleaner we denote R^n as V.

The following definition is the obvious generalization of Definition A.7 given in the appendix on page 745.

Definition 3.197. A function

$$\mathcal{A} : V^m \to V^r$$ [V]

is *m-linear*, or more generically *multi-linear*, if it is linear in each argument. That is for all integers $1 \leq i \leq m$, $\mathcal{A}$ satisfies

$$\mathcal{A}(\mathbf{v}_1, \ldots, a\mathbf{v}_i + b\mathbf{v}_i', \ldots, \mathbf{v}_m) = a\mathcal{A}(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_m) + b\mathcal{A}(\mathbf{v}_1, \ldots, \mathbf{v}_i', \ldots, \mathbf{v}_m).$$

Such a function is *alternating* (or anti-symmetric) if it is equal to $\mathbf{0}$ whenever two arguments are the same:

$$\mathcal{A}(\mathbf{v}_1, \ldots, \mathbf{v}, \ldots, \mathbf{v}, \ldots, \mathbf{v}_m) = \mathbf{0}.$$

For example the dot product, cross product, complex multiplication, and quaternonic multiplication are all 2-linear (simply called bilinear), but of these, only the cross product is alternating. The determinant of a 3 by 3 matrix is 3-linear (simply called trilinear) and alternating.

PROPOSITION 3.198. *Let $\mathcal{A}$ be an alternating m linear function. Then, for any m-tuple $(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ of vectors and any function $\sigma\colon \overline{m} \to \overline{m}$, we have*

$$\mathcal{A}(\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(m)}) = \varepsilon(\sigma) \cdot \mathcal{A}(\mathbf{v}_1, \ldots \mathbf{v}_m).$$

PROOF. (**Sketch**) If σ is not a permutation, then by the pigeon hole principle, for some $i \neq j \in \overline{n}$, $\sigma(i) = \sigma(j)$. Therefore, two of the arguments on the LHS of what we're trying to prove are equal so the whole thing is zero, since $\mathcal{A}$ is alternating. Moreover, in this case, by definition $\varepsilon(\sigma) = 0$ and therefore the RHS also equals zero and the equation checks in this case.

On the other hand, assume that σ is a permutation. The exact same proof as one would need to prove Exercise A.7 shows that $\mathcal{A}$ switches sign whenever two of its arguments switch places. That is

$$\mathcal{A}(\mathbf{v}_1, \ldots, \mathbf{v}, \ldots, \mathbf{w}, \ldots, \mathbf{v}_m) = -\mathcal{A}(\mathbf{v}_1, \ldots, \mathbf{w}, \ldots, \mathbf{v}, \ldots, \mathbf{v}_m).$$

Therefore, for every crossing of arguments an addition sign change occurs. But this is exactly how the sign of a permutation was defined in the first place. ■

To expand things in, say, a trilinear function you would expand things in exactly the same way you would if you had the product of three polynomials, for instance consider how you would expand the following

$$(y + 5z) \cdot (x + y + 2z) \cdot (7x + y)$$

without switching any of the variables. This would be the same as you expanding

$$\mathcal{A}(\mathbf{e}_2 + 5\mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3, 7\mathbf{e}_1 + \mathbf{e}_2).$$

Indeed, to do this out we see that

$$\begin{aligned}
&\mathcal{A}(\mathbf{e}_2+5\mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3, 7\mathbf{e}_1 + \mathbf{e}_2) \\
&\quad = \mathcal{A}(\mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3, 7\mathbf{e}_1 + \mathbf{e}_2) + \mathcal{A}(5\mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3, 7\mathbf{e}_1 + \mathbf{e}_2) \\
&\quad = \mathcal{A}(\mathbf{e}_2, \mathbf{e}_1, 7\mathbf{e}_1 + \mathbf{e}_2) + \mathcal{A}(\mathbf{e}_2, \mathbf{e}_2, 7\mathbf{e}_1 + \mathbf{e}_2) + \mathcal{A}(\mathbf{e}_2, 2\mathbf{e}_3, 7\mathbf{e}_1 + \mathbf{e}_2) \\
&\quad\ + \mathcal{A}(5\mathbf{e}_3, \mathbf{e}_1, 7\mathbf{e}_1 + \mathbf{e}_2) + \mathcal{A}(5\mathbf{e}_3, \mathbf{e}_2, 7\mathbf{e}_1 + \mathbf{e}_2) + \mathcal{A}(5\mathbf{e}_3, 2\mathbf{e}_3, 7\mathbf{e}_1 + \mathbf{e}_2) \\
&\quad = \mathcal{A}(\mathbf{e}_2, \mathbf{e}_1, 7\mathbf{e}_1) + \mathcal{A}(\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_2) \\
&\quad\ + \mathcal{A}(\mathbf{e}_2, \mathbf{e}_2, 7\mathbf{e}_1) + \mathcal{A}(\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_2) \\
&\quad\ + \mathcal{A}(\mathbf{e}_2, 2\mathbf{e}_3, 7\,\mathbf{e}_1) + \mathcal{A}(\mathbf{e}_2, 2\,\mathbf{e}_3, \mathbf{e}_2) \\
&\quad\ + \mathcal{A}(5\,\mathbf{e}_3, \mathbf{e}_1, 7\,\mathbf{e}_1) + \mathcal{A}(5\,\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2) \\
&\quad\ + \mathcal{A}(5\,\mathbf{e}_3, \mathbf{e}_2, 7\,\mathbf{e}_1) + \mathcal{A}(5\,\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_2) \\
&\quad\ + \mathcal{A}(5\,\mathbf{e}_3, 2\,\mathbf{e}_3, 7\,\mathbf{e}_1) + \mathcal{A}(5\,\mathbf{e}_3, 2\,\mathbf{e}_3, \mathbf{e}_2) \\
&\quad = 7\,\mathcal{A}(\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_1) + \mathcal{A}(\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_2) \\
&\quad\ + 7\,\mathcal{A}(\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_1) + \mathcal{A}(\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_2) \\
&\quad\ + 14\,\mathcal{A}(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1) + 2\,\mathcal{A}(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_2) \\
&\quad\ + 35\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_1) + 5\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2) \\
&\quad\ + 35\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1) + 5\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_2) \\
&\quad\ + 70\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_3, \mathbf{e}_1) + 10\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_3, \mathbf{e}_2).
\end{aligned} \tag{45}$$

Indeed, the polynomial just mentioned, $(y + 5z) \cdot (x + y + 2z) \cdot (7x + y)$, expands as

$$\begin{aligned}
(y + 5z) \cdot (x + y + 2z) \cdot (7x + y) = \quad & 7yx^2 + yxy + 7y^2x + y^3 \\
& + 14yzx + 2zy + 35zx^2 + 5zxy \\
& + 35zyx + 5zy^2 + 70z^2x + 10z^2y.
\end{aligned}$$

Now if we knew what $\mathcal{A}$ did to triples of basis elements we'd be able to calculate the numerical value of $\mathcal{A}$. This is sort-of analogous to multi-variable polynomials. We have been able to break down what $\mathcal{A}$ does to an (essentially) arbitrary triple of vectors to a sum of what $\mathcal{A}$ does to triples of standard basis vectors. At this point we can't really go any further without any further information. We give two possible examples of what we could know and show how to use the given information to finish the calculation.

- For instance, if we happened to know (I'm making this up)

$$\mathcal{A}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) = i \cdot j \cdot k$$

then I'd be able to complete my calculation started in (45):

$$\begin{aligned}
\mathcal{A}(\mathbf{e}_2 + 5\mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3, 7\mathbf{e}_1 + \mathbf{e}_2) &= \mathcal{A}(\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_1) + \mathcal{A}(\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_2) \\
&\quad + 7\,\mathcal{A}(\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_1) + \mathcal{A}(\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_2) \\
&\quad + 14\,\mathcal{A}(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1) + 2\,\mathcal{A}(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_2) \\
&\quad + 35\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_1) + 5\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2) \\
&\quad + 35\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1) + 5\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_2) \\
&\quad + 70\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_3, \mathbf{e}_1) + 10\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_3, \mathbf{e}_2) \\
&= 2 \cdot 1 \cdot 1 + 2 \cdot 1 \cdot 2 \\
&\quad + 7 \cdot 2 \cdot 2 \cdot 1 + 2 \cdot 2 \cdot 2 \\
&\quad + 14 \cdot 2 \cdot 3 \cdot 1 + 2 \cdot 2 \cdot 3 \cdot 2 \\
&\quad + 35 \cdot 3 \cdot 1 \cdot 1 + 5 \cdot 3 \cdot 1 \cdot 2 \\
&\quad + 35 \cdot 3 \cdot 2 \cdot 1 + 5 \cdot 3 \cdot 2 \cdot 2 \\
&\quad + 70 \cdot 3 \cdot 3 \cdot 1 + 10 \cdot 3 \cdot 3 \cdot 2 \\
&= 2 + 4 + 28 + 8 + 84 + 24 + 105 + 30 + 210 + 60 + 210 + 60 + 630 + 180 \\
&= 1635.
\end{aligned}$$

- If, on the other hand, if $\mathcal{A}$ happened to be alternating, then the last line of the general calculation would simplify dramatically since we can now cross off any terms which have any arguments which repeat:

$$\begin{aligned}
&7\,\mathcal{A}(\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_1) + \mathcal{A}(\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_2) + 7\,\mathcal{A}(\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_1) + \mathcal{A}(\mathbf{e}_2, \mathbf{e}_2, \mathbf{e}_2) \\
&\qquad + 14\,\mathcal{A}(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1) + 2\,\mathcal{A}(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_2) + 35\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_1) + 5\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2) \\
&\qquad\qquad + 35\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1) + 5\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_2) + 70\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_3, \mathbf{e}_1) + 10\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_3, \mathbf{e}_2)
\end{aligned}$$

which only leaves

$$14\,\mathcal{A}(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1) + 5\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2) + 35\,\mathcal{A}(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1).$$

Lastly, using the proposition, Proposition 3.198 coming up on page 182 (it says when we switch the arguments of an alternating multilinear functions we must tack on a sign change whenever we do this via an odd permutation), this simplifies further to

$$14\,\mathcal{A}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) + 5\,\mathcal{A}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) - 35\,\mathcal{A}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = -16\,\mathcal{A}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3).$$

If we don't know what $\mathcal{A}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ happened to be, we couldn't go any further. But the second we had a value for this, say $\mathcal{A}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = 1$ (as it does when $\mathcal{A} = \det$), then we'd have a final result of -16.

This kind of calculation shows that any multilinear map is uniquely determined by what it does to tuples of standard basis vectors. It also shows that for alternating multilinear maps, we only need to know what the map does to ordered tuples. This is analogous to the fact that we know the value of any polynomial $P(x_1, \ldots, x_n)$ the second we know what to plug in for the variables x_i's: in a sense, it is no longer a polynomial, it is now a constant.

More specifically, for $\mathbf{B} := (\mathbf{b}_1 \mid \cdots \mid \mathbf{b}_n)$ (where $(\mathbf{b}_1, \ldots, \mathbf{b}_n) \in V^n$) and an m by n matrix $\mathbf{A}$ with $\mathbf{A}(i, j) := a_{ij}$, then $\mathbf{W} := (\mathbf{w}_1 \mid \cdots \mid \mathbf{w}_m) = \mathbf{B} \cdot \mathbf{A}$, where

$$\mathbf{w}_1 := a_{11}\mathbf{b}_1 + \cdots + a_{1n}\mathbf{b}_n \tag{46}$$

$$\vdots \tag{47}$$

$$\mathbf{w}_m := a_{m1}\mathbf{b}_1 + \cdots + a_{mn}\mathbf{b}_n \tag{48}$$

in terms of what happens to sequences of $\mathbf{b}_i$ under $\mathcal{A}$. Indeed,

$$\begin{aligned}
\mathcal{A}(\mathbf{w}_1, \cdots, \mathbf{w}_m) &= \mathcal{A}(a_{11}\mathbf{b}_1 + \cdots + a_{1n}\mathbf{b}_n, a_{21}\mathbf{b}_1 + \cdots + a_{2n}\mathbf{b}_n, \cdots, a_{m1}\mathbf{b}_1 + \cdots + a_{mn}\mathbf{b}_n) \\
&= \sum_{i_1=1}^{n} a_{1i_1}\mathcal{A}(\mathbf{b}_{i_1}, a_{21}\mathbf{b}_1 + \cdots + a_{2n}\mathbf{b}_n, \cdots, a_{m1}\mathbf{b}_1 + \cdots + a_{mn}\mathbf{b}_n) \\
&= \sum_{i_1=1}^{n} a_{1i_1}\left(\sum_{i_2=1}^{n} a_{2i_2}\mathcal{A}(\mathbf{b}_{i_1}, \mathbf{b}_{i_2}, \cdots, a_{m1}\mathbf{b}_1 + \cdots + a_{mn}\mathbf{b}_n)\right) \\
&= \sum_{i_1=1}^{m} a_{1i_1}\left(\sum_{i_2=1}^{m} a_{2i_2}\cdots\left(\sum_{i_m=1}^{m} a_{mi_m}\mathcal{A}(\mathbf{b}_{i_1}, \mathbf{b}_{i_2}, \cdots, \mathbf{b}_{i_m})\right)\right) \\
&= \sum_{(i_1,\ldots,i_m)\in\overline{n}^m} a_{1i_1}\cdots a_{mi_m} \cdot \mathcal{A}(\mathbf{b}_{i_1}, \cdots, \mathbf{b}_{i_m}).
\end{aligned}$$

Since if we take the $\mathbf{b}_i$ to be $\mathbf{e}_i$, then every such $(\mathbf{w}_1, \ldots, \mathbf{w}_m)$ can be expressed as in Equation (46) proving:

THEOREM 3.199. *An m–linear function $\mathcal{A}$ is uniquely determined by what it does to the n^m m–tuples of the n standard basis vectors*

$$\mathbf{e}_1, \ \ldots, \mathbf{e}_n.$$

PROOF. ∎

Let's examine the case that $n = m$ and $\mathcal{A}$ is assumed alternating. Here, for any n-tuple of numbers, $(i_1, \cdots, i_n)$. Then, by Proposition 3.198, we have

$$\mathcal{A}(\mathbf{b}_{i_1}, \cdots, \mathbf{b}_{i_n}) = \varepsilon\begin{pmatrix} 1 & \cdots & n \\ i_1 & \cdots & i_n \end{pmatrix} \cdot \mathcal{A}(\mathbf{b}_1, \cdots, \mathbf{b}_n)$$

where σ is the function

$$\sigma := \begin{pmatrix} 1 & \cdots & n \\ i_1 & \cdots & i_n \end{pmatrix}.$$

Therefore

$$\begin{aligned}\mathcal{A}(\mathbf{w}_1,\cdots,\mathbf{w}_n) &= \sum_{(i_1,\ldots,i_n)\in\overline{n}^n} a_{1i_1}\cdots a_{ni_n}\,\varepsilon\begin{pmatrix}1 & \cdots & n\\ i_1 & \cdots & i_n\end{pmatrix}\cdot\mathcal{A}(\mathbf{b}_1,\cdots,\mathbf{b}_n)\\ &= \mathcal{A}(\mathbf{b}_1,\cdots,\mathbf{b}_n)\cdot\left(\sum_{(i_1,\ldots,i_n)\in\overline{n}^n}\varepsilon\begin{pmatrix}1 & \cdots & n\\ i_1 & \cdots & i_n\end{pmatrix}\cdot a_{1i_1}\cdots a_{ni_n}\right).\end{aligned}$$

We can jazz this up even further since

$$\varepsilon\begin{pmatrix}1 & \cdots & n\\ i_1 & \cdots & i_n\end{pmatrix} = 0$$

whenever $(i_1,\ldots,i_n)\in\overline{n}^n\setminus\overline{n}^{*n}$. Therefore, identifying $\mathcal{S}_n$ with $\overline{n}^{*n}$, our sum is over the permutations and

$$\mathcal{A}(\mathbf{w}_1,\ldots,\mathbf{w}_n) = \mathcal{A}(\mathbf{b}_1,\ldots,\mathbf{b}_n)\cdot\left(\sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)\cdot a_{1\sigma(1)}\cdot\cdots\cdot a_{n\sigma(n)}\right). \tag{49}$$

We can think of the n-multilinear function $\mathcal{A}$ as having its domain $\mathbf{M}_n(R)$ instead of V^n via

$$\mathcal{A}(\mathbf{w}_1\mid\cdots\mid\mathbf{w}_n)\overset{\text{def}}{=}\mathcal{A}(\mathbf{w}_1,\ldots,\mathbf{w}_n).$$

In this way we've proven that

$$\mathcal{A}(\mathbf{W}) = \mathcal{A}(\mathbf{B})\cdot\left(\sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)\cdot\mathbf{A}(1,\sigma(1))\cdot\cdots\cdot\mathbf{A}(n,\sigma(n))\right). \tag{50}$$

whenever $\mathbf{W}=\mathbf{B}\cdot\mathbf{A}$. Taking $\mathbf{B}=\mathbf{I}_n$, we have that

$$\mathcal{A}(\mathbf{A}) = \mathcal{A}(\mathbf{I}_n)\cdot\left(\sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)\cdot\mathbf{A}(1,\sigma(1))\cdot\cdots\cdot\mathbf{A}(n,\sigma(n))\right). \tag{51}$$

Now the summed quantity

$$\sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)\cdot\mathbf{A}(1,\sigma(1))\cdot\cdots\cdot\mathbf{A}(n,\sigma(n))$$

has **nothing** to do with $\mathcal{A}$, it only has to do with how with the matrix $\mathbf{A}$. We temporarily denote this quantity as $\mathrm{D}(\mathbf{A})$, i.e.

$$\mathrm{D}(\mathbf{A})\overset{\text{def}}{=}\sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)\cdot\mathbf{A}(1,\sigma(1))\cdot\cdots\cdot\mathbf{A}(n,\sigma(n)).$$

I say *temporarily* since we'll show below that $\mathrm{D}(\mathbf{A})=\det\mathbf{A}$. For now we have just by plugging the definition of D into (51).

$$\mathcal{A}(\mathbf{A}) = \mathcal{A}(\mathbf{I}_n)\cdot\mathrm{D}(\mathbf{A}). \tag{52}$$

Again, since $\mathrm{D}(\mathbf{A})$ is independent of $\mathcal{A}$, we have

PROPOSITION 3.200. *Any n-linear, alternating function, $\mathcal{A}$, is uniquely determined by the value $\mathcal{A}(\mathbf{I}_n)$.* ■

Note that this proposition should have been predictable from the example above where we were able that

$$\mathcal{A}(\mathbf{e}_2 + 5\mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_3, 7\mathbf{e}_1 + \mathbf{e}_2) = -16\mathcal{A}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3).$$

Therefore the second we are given what $\mathcal{A}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ is we just plug it in and we are done.

EXERCISE 3.75. Suppose that $\mathcal{A}$ is a trilinear function.

SOLUTION

(1) Expand

$$\mathcal{A}(2\mathbf{e}_1 + \mathbf{e}_3 + \mathbf{e}_4,\ \mathbf{e}_1 + \mathbf{e}_4,\ 2\mathbf{e}_2 + \mathbf{e}_4)$$

as much as possible.

(2) Repeat the previous part with the added assumption that $\mathcal{A}$ is alternating (you won't be able to get it down to a single term like I was with the example above. The reason why is I have 3 linear function in a 4 dimensional situation.)

(3) Now assume that $\mathcal{A}$ takes on the following values

$$\mathcal{A}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) = i + j + k, \quad 1 \leq i, j, k \leq 4.$$ [VI]

(Note that this implies that $\mathcal{A}$ is not alternating and so the previous part does not apply.) Get a numerical value for what you started in the first part.

NOTATION 3.201. (**Axiomatic**) Det, is the unique alternating, n-linear function

$$\text{Det}: \mathbf{M}_n(R) \to R$$

with

$$\text{Det}(\mathbf{I}_n) = 1.$$

For the next theorem we use the following notations:

NOTATION 3.202. For an n-vector ($n > 1$), $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ define $\overline{\mathbf{v}}$ to be the $n-1$ vector given by the last $n-1$ components of $\mathbf{v}$:

$$\overline{\mathbf{v}} \stackrel{\text{def}}{=} \begin{bmatrix} v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

[VI] So for instance $\mathcal{A}(\mathbf{e}_1, \mathbf{e}_1, \mathbf{e}_1) = 1 + 1 + 1 = 3$ and $\mathcal{A}(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1) = 3 + 2 + 1 = 6$.

Define $\mathbf{v}'$ as the first component of $\mathbf{v}$, that is $\mathbf{v}' = v_1$.

Clearly the maps $\mathbf{v} \mapsto \overline{\mathbf{v}}$ and $\mathbf{v} \mapsto \mathbf{v}'$ are linear. With this notation, we have that

$$\det((\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n)) = \sum_{i=1}^{n} \mathbf{v}_i' \cdot \det\left((\overline{\mathbf{v}}_1 \mid \cdots \mid \widehat{\overline{\mathbf{v}}_i} \mid \cdots \mid \overline{\mathbf{v}}_n)\right)$$

where the $\widehat{\ }$ appearing over a vector means that the indicated vector is to be omitted from the computation.

THEOREM 3.203.

$$\mathrm{Det} = \det.$$

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} = \begin{pmatrix} \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & 1 & \cdots & 0 & 0 \\ \blacksquare & \vdots & & \vdots & \vdots \\ \blacksquare & 0 & \cdots & 1 & 0 \\ \blacksquare & 0 & \cdots & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} = \mathbf{I}_{n-1}$$

Figure 3.6: $\mathbf{I}_n^{(1,1)} = \mathbf{I}_{n-1}$

PROOF. We need to prove

- $\det \mathbf{I}_n = 1$;
- det is n-linear;
- det is alternating.

For an n-vector $\mathbf{v}$ we define $\overline{\mathbf{v}}$ to be the $n-1$-vector given by the last $n-1$ components of $\mathbf{v}$ and $\mathbf{v}'$ as the first component of $\mathbf{v}$.

(1) Since det of a 1 by 1 matrix (r) is r, we know that $\det(1) = 1$. Now proceeding by induction, assuming that $\det \mathbf{I}_{n-1} = 1$, we have

$$\det \mathbf{I}_n = \sum_{k=1}^{n} (-1)^{1+k} \mathbf{I}_n(1,k) \cdot \det\left(\mathbf{I}_n{}^{(1,k)}\right) = (-1)^{\overset{1}{\shortparallel}+1} \mathbf{I}_n(\overset{1}{\shortparallel}, 1) \cdot \det\left(\mathbf{I}_n{}^{(1,1)}\right) = \det\left(\mathbf{I}_n^{(1,1)}\right).$$

But $\mathbf{I}_n^{(1,1)}$ is clearly just $\mathbf{I}_{n-1}$, see Figure 3.6. Therefore, using the induction hypothesis:

$$\det \mathbf{I}_n = \det \mathbf{I}_n^{(1,1)} = \det \mathbf{I}_{n-1} = 1,$$

and we are done.

(2) The determinant of a 1 by 1 matrix, being just the identity function, is 1-linear. Assuming the determinant of $n-1$ by $n-1$ matrix is $n-1$ linear. We can easily prove that this is also true for the n by n case simply by using the linearity of the functions $\mathbf{v} \mapsto \overline{\mathbf{v}}$ and $\mathbf{v} \mapsto \mathbf{v}'$.

(3) The first case we need to check here is the 2 by 2 case since the 1 by 1 case is vacuously alternating. Indeed, this can be checked directly by the formula. Now assume that the $n-1$ by $n-1$ det is alternating. Suppose $\mathbf{v}_i = \mathbf{v}_j$ for distinct i and j which we may as well assume $i < j$. We need to prove that

$$\det((\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n)) = 0.$$

Since we also have $\mathbf{v}_i = \mathbf{v}_j \Rightarrow \overline{\mathbf{v}}_i = \overline{\mathbf{v}}_j$, every term in the sum

$$\det((\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n)) = \sum_{k=1}^{n} (-1)^{1+k} \mathbf{v}'_k \cdot \det\left((\overline{\mathbf{v}}_1 \mid \cdots \mid \widehat{\overline{\mathbf{v}}_k} \mid \cdots \mid \overline{\mathbf{v}}_n)\right)$$

is 0 unless k is either i or j. This is because for other terms where k is neither i nor j, both $\overline{\mathbf{v}}_i$ and $\overline{\mathbf{v}}_j$ show up as arguments and hence causing the summand to equal 0 by the inductive hypothesis. We therefore have that

$$\begin{aligned}
\det(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n) &= (-1)^{i+1}\mathbf{v}'_i \cdot \det\left(\overline{\mathbf{v}}_1 \mid \cdots \mid \widehat{\overline{\mathbf{v}}_i} \mid \cdots \mid \overline{\mathbf{v}}_j \mid \cdots \mid \overline{\mathbf{v}}_n\right) + (-1)^{j+1}\mathbf{v}'_j \cdot \det\left(\overline{\mathbf{v}}_1 \mid \cdots \mid \overline{\mathbf{v}}_i \mid \cdots \mid \widehat{\overline{\mathbf{v}}_j} \mid \cdots \mid \overline{\mathbf{v}}_n\right) \\
&= (-1)^{i+1}\mathbf{v}'_i \cdot \det\left(\overline{\mathbf{v}}_1 \mid \cdots \mid \widehat{\overline{\mathbf{v}}_i} \mid \cdots \mid \overline{\mathbf{v}}_i \mid \cdots \mid \overline{\mathbf{v}}_n\right) + (-1)^{j+1}\mathbf{v}'_i \cdot \det\left(\overline{\mathbf{v}}_1 \mid \cdots \mid \overline{\mathbf{v}}_i \mid \cdots \mid \widehat{\overline{\mathbf{v}}_i} \mid \cdots \mid \overline{\mathbf{v}}_n\right) \\
&= \mathbf{v}'_i \cdot \left((-1)^{i+1}\det\left(\overline{\mathbf{v}}_1 \mid \cdots \mid \widehat{\overline{\mathbf{v}}_i} \mid \cdots \mid \overline{\mathbf{v}}_i \mid \cdots \mid \overline{\mathbf{v}}_n\right) + (-1)^{j+1}\det\left(\overline{\mathbf{v}}_1 \mid \cdots \mid \overline{\mathbf{v}}_i \mid \cdots \mid \widehat{\overline{\mathbf{v}}_i} \mid \cdots \mid \overline{\mathbf{v}}_n\right)\right). \qquad (53)
\end{aligned}$$

The lists involved in these two terms are (I overset what coordinate the given vectors lives in)

$$(\overset{1}{\overline{\mathbf{v}}_1}, \ldots, \overset{i-1}{\overline{\mathbf{v}}_{i-1}}, \overset{i}{\overline{\mathbf{v}}_i}, \overset{i+1}{\overline{\mathbf{v}}_{i+1}}, \ldots, \overset{j-1}{\overline{\mathbf{v}}_{j-1}}, \overset{j}{\overline{\mathbf{v}}_{j+1}}, \ldots, \overset{n-1}{\overline{\mathbf{v}}_n}) \text{ and } (\overset{1}{\overline{\mathbf{v}}_1}, \ldots, \overset{i-1}{\overline{\mathbf{v}}_{i-1}}, \overset{i}{\overline{\mathbf{v}}_{i+1}}, \overset{i+1}{\overline{\mathbf{v}}_{i+2}}, \ldots, \overset{j-2}{\overline{\mathbf{v}}_{j-1}}, \overset{j-1}{\overline{\mathbf{v}}_i = \overline{\mathbf{v}}_j}, \overset{j}{\overline{\mathbf{v}}_{j+1}}, \ldots, \overset{n-1}{\overline{\mathbf{v}}_n}).$$

They differ by a permutation

$$\sigma := \begin{pmatrix} 1 & \cdots & i-1 & i & i+1 & \cdots j-1 & j & j+1 & \cdots n \\ 1 & \cdots & i-1 & j-1 & i & \cdots j-2 & j & j+1 & \cdots n \end{pmatrix}.$$

Inductively, det on $n-1$ vectors is alternating. As such we have that

$$\det\left(\overline{\mathbf{v}}_1 \mid \cdots \mid \overline{\mathbf{v}}_i \mid \cdots \mid \widehat{\overline{\mathbf{v}}_i} \mid \cdots \mid \overline{\mathbf{v}}_n\right) = \varepsilon(\sigma) \cdot \det\left(\overline{\mathbf{v}}_1 \mid \cdots \mid \widehat{\overline{\mathbf{v}}_i} \mid \cdots \mid \overline{\mathbf{v}}_i \mid \cdots \mid \overline{\mathbf{v}}_n\right). \qquad (54)$$

It remains to calculate $\varepsilon(\sigma)$. This is easy: σ has the letter i crossing over the $j-i-1$ letters between i and $j-1$ and no other crossings. This means the crossing number of σ is $j-i-1$, and so the sign σ is $(-1)^{j-i-1}$. Therefore, plugging into Equation (54), we see

$$\det\left(\overline{\mathbf{v}}_1 \mid \cdots \mid \overline{\mathbf{v}}_i \mid \cdots \mid \widehat{\overline{\mathbf{v}}_i} \mid \cdots \mid \overline{\mathbf{v}}_n\right) = (-1)^{j-i-1} \cdot \det\left(\overline{\mathbf{v}}_1 \mid \cdots \mid \widehat{\overline{\mathbf{v}}_i} \mid \cdots \mid \overline{\mathbf{v}}_i \mid \cdots \mid \overline{\mathbf{v}}_n\right).$$

Since $\mathbf{v}_i = \mathbf{v}_j$ we also have that $\mathbf{v}'_i = \mathbf{v}'_j$. Therefore the two terms in (53) are the same but with the second having sign

$$(-1)^{j+1} \cdot (-1)^{j-i-1} = (-1)^{2j} \cdot (-1)^2 \cdot (-1)^i = (-1)^i$$

and the first having sign $(-1)^{i+1} = -(-1)^i$, which is the negative of the second. Therefore that sum must be zero. ∎

This means that, using Equation (52), that

$$\det(\mathbf{A}) = \det(\mathbf{I}_n) \cdot \mathrm{D}(\mathbf{A}) = \mathrm{D}(\mathbf{A}).$$

Plugging this back into Corollary 3.200 proves

THEOREM 3.204. *For any n-alternating function $\mathcal{A}$, we have*

$$\mathcal{A}(\mathbf{B} \cdot \mathbf{A}) = \mathcal{A}(\mathbf{B}) \cdot \det(\mathbf{A}). \qquad (55)$$

In particular, taking $\mathcal{A} = \det$ proves

THEOREM 3.205.

(1) *For any*

$$(\mathbf{v}_1, \ldots, \mathbf{v}_n) \in V^n$$

and any $\mathbf{A} \in \mathbf{M}_n(R)$ *we have*

$$\det(\mathbf{A}(\mathbf{v}_1), \ldots, \mathbf{A}(\mathbf{v}_n)) = \det \mathbf{A} \cdot \det(\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

(2) *For every* $\mathbf{A}, \mathbf{B} \in \mathbf{M}_n(R)$ *we have*

$$\det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{A}) \cdot \det(\mathbf{B}).$$

■

Let's see how to do calculations using the axiomatic definition:

Using the axiomatic definition to calculate the determinant

(1) Break up the argument of the determinant (given as a n-tuple of n-vectors) by using the n-linearity of the determinant to break up the argument as a linear combination of arguments of the form

$$\det(\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)})$$

for some permutation (or just a function that needn't be one-to-one).

(2) Use the proposition, Proposition 3.198, to rewrite each

$$\det(\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)})$$

as

$$\varepsilon(\sigma) \cdot \det(\mathbf{e}_1, \ldots, \mathbf{e}_n).$$

(3) Use the last part of the axiomatic version to replace each of the $\det(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ by 1.

(4) Add what's left to get the determinant.

EXAMPLE 3.206. We use the procedure to calculate the determinant of the matrix

$$\begin{pmatrix} 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 \end{pmatrix}.$$

In this calculation we make repeated use of the fact that for an alternating multilinear function $\mathcal{A}$ we have

$$\mathcal{A}(\ldots, \mathbf{v}, \ldots, \mathbf{v} + \mathbf{w}, \ldots) = \cancel{\mathcal{A}(\ldots, \mathbf{v}, \ldots, \mathbf{v}, \ldots)} + \mathcal{A}(\ldots, \mathbf{v}, \ldots, \mathbf{w}, \ldots) = \mathcal{A}(\ldots, \mathbf{v}, \ldots, \mathbf{w}, \ldots)$$

where we could cross out because the argument, $\mathbf{v}$, was repeated. Let's get down to the calcuation:

$$\begin{aligned}
&\det(2\mathbf{e}_1 + \mathbf{e}_3,\ \mathbf{e}_1 + \mathbf{e}_4, 2\mathbf{e}_2 + \mathbf{e}_4,\ \mathbf{e}_1 + 2\mathbf{e}_3) \\
&\qquad = 2\det(\mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_4, 2\mathbf{e}_2 + \mathbf{e}_4, \mathbf{e}_1 + 2\mathbf{e}_3) + \det(\mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_4, 2\mathbf{e}_2 + \mathbf{e}_4, \mathbf{e}_1 + 2\mathbf{e}_3) \\
&\qquad = 4\det(\mathbf{e}_1, \mathbf{e}_4, 2\mathbf{e}_2 + \mathbf{e}_4, \mathbf{e}_3) + \det(\mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_4, 2\mathbf{e}_2 + \mathbf{e}_4, \mathbf{e}_1) \\
&\qquad = 8\det(\mathbf{e}_1, \mathbf{e}_4, \mathbf{e}_2, \mathbf{e}_3) + \det(\mathbf{e}_3, \mathbf{e}_4, 2\mathbf{e}_2 + \mathbf{e}_4, \mathbf{e}_1) \\
&\qquad = 8\det(\mathbf{e}_1, \mathbf{e}_4, \mathbf{e}_2, \mathbf{e}_3) + 2\det(\mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_2, \mathbf{e}_1) \\
&\qquad = 8\varepsilon\begin{pmatrix}1&2&3&4\\1&4&2&3\end{pmatrix} + 2\varepsilon\begin{pmatrix}1&2&3&4\\3&4&2&1\end{pmatrix} \\
&\qquad = 8\cdot(+1) + 2\cdot(-1) \\
&\qquad = 6.
\end{aligned}$$

EXERCISE 3.76. Using only the inductive definition, prove that the determinant is multilinear.
Hint! The is a straight forward induction, but you need to make two cases based on whether you are dealing with the first row or not.

SOLUTION

This exercise is harder than the previous one.

EXERCISE 3.77. Using only the inductive definition, prove that the determinant is alternating.
Hint! Again use induction. You can reduce everything down to proving the result when the first and second rows happen to be the same. Then you can finish the proof, but it does take a lot of concentration. Work some concrete examples to gain confidence!

SOLUTION

EXERCISE 3.78. Use the axiomatic definition to calculate the determinant of the matrices

$$\begin{pmatrix}1&0&1&1\\0&1&2&0\\0&2&0&1\\1&0&2&0\end{pmatrix} \text{ and } \begin{pmatrix}[1]_3&[2]_3&[0]_3\\[2]_3&[2]_3&[2]_3\\[2]_3&[2]_3&[1]_3\end{pmatrix}.$$

SOLUTION

Recall that an element $r \in R$ is a *zero divisor* if there exists a non-zero $s \in R$ so that $s \cdot r = 0$.

EXERCISE 3.79. Show that

$$\det(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n)$$

is a zero-divisor whenever $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are linearly dependent.

SOLUTION

In fact, the converse of the statement of this exercise is also true (whenever R contains no zero-divisors). The proof given in any linear algebra course, which is usually given only for $R = \mathbb{R}$, carries over verbatim to the current situation. (In fact we could prove this fact after just a slight detour.) To see that we need to require that R not contain zero-divisors consider the 1-vector

$$\mathbf{v} = ([2]_4).$$

Then since $[2]_4 \mathbf{v} = \mathbf{0}$, we would have that (by itself!) $\{[2]_4\}$ is linearly dependent (in the definition we are taking $n = 1$, $r_1 = [2]_4$ and $\mathbf{v} = \mathbf{v}_1 = [2]_4$), but the determinant of the matrix $([2]_4)$ is $[2]_4$, which does not equal 0 (i.e not equal to $[0]_4$).

3.5.3.3 Method Three: Formulaic

We now focus on the formula for the determinant which allows us to compute the determinant directly. Due to the complexity of the formula (which grows as $n!$ in length), I'm sure you'll be convinced that it's probably a good thing that you have haven't seen it until now.

By our earlier calculation, we also get our third definition, which we deal with in this section:

Theorem 3.207. *(**Formulaic Version**) For an n by n matrix* $\mathbf{A}$*, the determinant is*

$$\det \mathbf{A} = \sum_{\sigma \in S_n} \varepsilon(\sigma) \cdot \mathbf{A}(1, \sigma(1)) \cdots \mathbf{A}(n, \sigma(n)).$$

For instance this says the determinant of the general 4 by 4 matrix

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix}$$

is given by the formula (which of course has $4! = 24$ terms)

$$\begin{aligned} \sum_{\sigma \in S_4} \varepsilon(\sigma) \cdot A_{1,\sigma(1)} A_{2,\sigma(2)} A_{3,\sigma(3)} A_{4,\sigma(4)} = \\ + A_{11}A_{22}A_{33}A_{44} - A_{11}A_{22}A_{34}A_{43} + A_{11}A_{32}A_{43}A_{24} - A_{11}A_{32}A_{23}A_{44} \\ + A_{11}A_{42}A_{23}A_{34} - A_{11}A_{42}A_{33}A_{24} - A_{21}A_{12}A_{33}A_{44} + A_{21}A_{12}A_{34}A_{43} \\ - A_{21}A_{32}A_{43}A_{14} + A_{21}A_{32}A_{13}A_{44} - A_{21}A_{42}A_{13}A_{34} + A_{21}A_{42}A_{33}A_{14} \\ + A_{31}A_{12}A_{23}A_{44} - A_{31}A_{12}A_{43}A_{24} + A_{31}A_{22}A_{43}A_{14} - A_{31}A_{22}A_{13}A_{44} \\ + A_{31}A_{42}A_{13}A_{24} - A_{31}A_{42}A_{23}A_{14} - A_{41}A_{12}A_{23}A_{34} + A_{41}A_{12}A_{33}A_{24} \\ - A_{41}A_{22}A_{33}A_{14} + A_{41}A_{22}A_{13}A_{34} - A_{41}A_{32}A_{13}A_{24} + A_{41}A_{32}A_{23}A_{14}. \end{aligned}$$

Exercise 3.80. Calculate the coefficient of the following in the formula for a general 5 by 5 matrix

(1) $A_{1,5}A_{2,3}A_{3,1}A_{4,2}A_{5,4}$

(2) $A_{1,3}A_{2,1}A_{3,4}A_{4,3}A_{5,2}$.

SOLUTION

3.5.3.4 Some Exercises to Compare

As I already pointed out, we won't actually show that the three definitions are the same. {Reference a source that does} We simply give a couple of exercises which allow you to verify this statement.

Exercise 3.81. Compute the determinant of

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 \end{pmatrix}$$

using all three methods. Which took you the least amount of time?

Solution

Exercise 3.82. Let $\mathbf{A} \in \mathbf{M}_n(\mathbb{Z})$ (i.e. and n by n matrix with integer coefficients). Define $[\mathbf{A}]_n \in \mathbf{M}_{[}(n)][\mathbb{Z}_n]$ as the matrix one obtains by taking all entries mod n. [VII] Prove that

$$[\det \mathbf{A}]_n = \det([\mathbf{A}]_n).$$

Solution

3.5.3.5 Theorems About the Determinant

For us the most important theorem about the determinant, in words, is the that the determinant of a product is the product of the determinants. In the case that our matrices have real coefficients, this is proven in nearly any elementary linear algebra textbook. We only sketch the proof here.

Corollary 3.208. *For any two matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *as in the theorem we have that*

(1) $\det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{B} \cdot \mathbf{A})$,

And if we moreover assume that $\mathbf{B}$ *is invertible, then*

(2) $\det(\mathbf{B}^{-1}) = \det(\mathbf{B})^{-1}$

(3) $\det(\mathbf{B} \cdot \mathbf{A} \cdot \mathbf{B}^{-1}) = \det(\mathbf{A})$.

Proof.

(1)

$$\begin{aligned} \det(\mathbf{A} \cdot \mathbf{B}) &= \det(\mathbf{A}) \cdot \det(\mathbf{B}) && \text{(By the theorem)} \\ &= \det(\mathbf{B}) \cdot \det(\mathbf{A}) && (R \text{ assumed commutative}) \\ &= \det(\mathbf{B} \cdot \mathbf{A}). && \text{(By the theorem)} \end{aligned}$$

(2) Since $\mathbf{B} \cdot \mathbf{B}^{-1} = \mathbf{I}_n$ we have that $\det(\mathbf{B}) \cdot \det(\mathbf{B}^{-1}) = \det \mathbf{I}_n$ by the theorem. But $\det \mathbf{I}_n$ is obviously just 1. Therefore

$$\begin{aligned} &\det(\mathbf{B}) \cdot \det(\mathbf{B}^{-1}) = 1 \\ \Rightarrow\ &\cancel{\det(\mathbf{B})^{-1}}\ \cancel{\det(\mathbf{B})} \cdot \det(\mathbf{B}^{-1}) = \det(\mathbf{B})^{-1} \\ \Rightarrow\ &\det(\mathbf{B}^{-1}) = \det(\mathbf{B})^{-1}. \end{aligned}$$

which was the desired result.

(3) Again, this uses the fact that we are assuming that R is commutative. Indeed,

$$\begin{aligned}
\det\left(\mathbf{B}\cdot\mathbf{A}\cdot\mathbf{B}^{-1}\right) &= \det(\mathbf{B})\cdot\det(\mathbf{A})\cdot\det\left(\mathbf{B}^{-1}\right) && \text{(Two uses of thm)}\\
&= \det(\mathbf{B})\cdot\det(\mathbf{B}^{-1})\cdot\det(\mathbf{A}) && (R \text{ is commutative})\\
&= 1\cdot\det\mathbf{A} && \text{(Previous part)}\\
&= \det\mathbf{A}.
\end{aligned}$$

as desired. ∎

3.5.4 The Transpose of a Matrix

Definition 3.209. Let $\mathbf{A}$ be a matrix with coefficients in some commutative ring. The transpose of $\mathbf{A}$, denoted $\mathbf{A}^T$ is the matrix one obtains from flipping A across its diagonal. More formally,

$$(\mathbf{A}^\top)(i,j) = \mathbf{A}(j,i).$$

Remark 3.210.

(1) Notice that if $\mathbf{A}$ is an n by m matrix, then $\mathbf{A}^\top$ is an m by n matrix.

(2) Notice that for vectors (thought of as 1 by n matrices) $\mathbf{v}$ and $\mathbf{w}$ we have

$$\mathbf{v}^\top\cdot\mathbf{w} = \mathbf{v}\bullet\mathbf{w},$$

the dot product of the two vectors.

(3) More generally, using Remark 3.173, we see that for any two m by n matrices written as columns $\mathbf{B} = (\mathbf{b}_1|\cdots|\mathbf{b}_n)$ and $\mathbf{A} = (\mathbf{a}_1|\cdots|\mathbf{a}_n)$, since the columns of $\mathbf{B}$ are the rows of $\mathbf{B}^\top$,

$$\mathbf{B}^\top\cdot\mathbf{A} = \begin{pmatrix} \mathbf{b}_1\bullet\mathbf{a}_1 & \cdots & \mathbf{b}_1\bullet\mathbf{a}_n \\ \vdots & \vdots & \vdots \\ \mathbf{b}_n\bullet\mathbf{a}_1 & \cdots & \mathbf{b}_n\bullet\mathbf{a}_n. \end{pmatrix}$$

In other words

$$\left(\mathbf{B}^\top\cdot\mathbf{A}\right)(i,j) = \mathbf{b}_i\bullet\mathbf{a}_j.$$

In particular, $(\mathbf{A}^\top\cdot\mathbf{A})(i,j) = \mathbf{a}_i\bullet\mathbf{a}_j$ which happens to equal $\mathbf{I}_n$ iff the columns of $\mathbf{A}$ form an orthonormal basis.

Example 3.211.

(1)

$$\begin{pmatrix} 1 & 0 & 2 & 2 \\ 0 & 2 & 2 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}^{\top} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 2 & 2 & 0 \\ 2 & 1 & 1 \end{pmatrix}$$

(2)

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 0 \\ 1 & 2 & 2 \end{pmatrix}^{\top} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 1 & 0 & 2 \end{pmatrix}$$

(3) For any positive integer, the transpose of the identity matrix, $\mathbf{I}_n$ is just the identity matrix itself. That is $(\mathbf{I}_n)^{\top} = \mathbf{I}_n$. More generally, for any diagonal matrix $\mathbf{D}$, meaning $\mathbf{D}(i, j) = 0$ whenever $i \neq j$ we also have that $\mathbf{D}^{\top} = \mathbf{D}$.

THEOREM 3.212. *For matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *we have* [VIII]

(1) $(\mathbf{A}^{\top})^{\top} = \mathbf{A}$.

(2) *Transpose in linear. That is*

- **(a)** $(\mathbf{A} + \mathbf{B})^{\top} = \mathbf{A}^{\top} + \mathbf{B}^{\top}$
- **(b)** $(c \cdot \mathbf{A})^{\top} = c \cdot \mathbf{A}^{\top}$

(3) $(\mathbf{A} \cdot \mathbf{B})^{\top} = \mathbf{B}^{\top} \cdot \mathbf{A}^{\top}$

(4) $(\mathbf{A}^{\top})^{-1} = \left(\mathbf{A}^{-1}\right)^{\top}$

(5) $\det(\mathbf{A}^{\top}) = \det(\mathbf{A})$.

PROOF.

(1) This should be intuitively obvious. When swap rows and columns and then you swap them again, you end up where you started.

$$(\mathbf{A}^{\top})^{\top}(i, j) = \mathbf{A}^{\top}(j, i) = \mathbf{A}(i, j).$$

(2) We have

$$\begin{aligned} (\mathbf{A} + \mathbf{B})^{\top}(i, j) &= (\mathbf{A} + \mathbf{B})(j, i) \\ &= \mathbf{A}(j, i) + \mathbf{B}(j, i) \\ &= \mathbf{A}^{\top}(i, j) + \mathbf{B}^{\top}(i, j). \end{aligned}$$

The proof that $^{\top}$ preserves scalar multiplication is just as easy.

[VIII] Of course we are making assumptions on the dimensions of $\mathbf{A}$ and $\mathbf{B}$ so that each of these statements actually make sense. For instance in the the second part we are assuming that both $\mathbf{A}$ and $\mathbf{B}$ have the same size and in the third part we are assuming that the number of rows of $\mathbf{B}$ is equal to the number of columns of $\mathbf{A}$.

(3) We have

$$\begin{aligned}(\mathbf{A}\cdot\mathbf{B})^\top(i,j) &= (\mathbf{A}\cdot\mathbf{B})(j,i)\\ &= \sum_{k=0}^{n}\mathbf{A}(j,k)\cdot\mathbf{B}(k,i)\\ &= \sum_{k=0}^{n}\mathbf{A}^\top(k,j)\cdot\mathbf{B}^\top(i,k)\\ &= \sum_{k=0}^{n}\mathbf{B}^\top(i,k)\cdot\mathbf{A}^\top(k,j) \quad \text{(Coefficients of } \mathbf{B} \text{ and } \mathbf{A} \text{ commute}^{\text{IX}}\text{)}\\ &= (\mathbf{B}^\top\cdot\mathbf{A}^\top)(i,j).\end{aligned}$$

Thus every coefficient of the matrix $(\mathbf{A}\cdot\mathbf{B})^\top$ is equal to the corresponding coefficient of the matrix $\mathbf{B}^\top\cdot\mathbf{A}^\top$ and, as such, the matrices themselves are equal.

(4) To show that $\left(\mathbf{A}^{-1}\right)^\top = (\mathbf{A}^\top)^{-1}$, we need to multiply the LHS by $\mathbf{A}^\top$ and see that we get the identity. Indeed,

$$\begin{aligned}\left(\mathbf{A}^{-1}\right)^\top\cdot\mathbf{A}^\top &= \left(\mathbf{A}\cdot\mathbf{A}^{-1}\right)^\top && \text{(By the previous part)}\\ &= (\mathbf{I}_n)^\top && (\mathbf{A}\,\mathbf{A}^{-1}=\mathbf{I}_n)\\ &= \mathbf{I}_n. && \text{(Trans of id is id)}\end{aligned}$$

(5) [X] Recall that the formulaic definition of the determinant says that

$$\det(\mathbf{A}) = \sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)\cdot\mathbf{A}(1,\sigma(1))\cdots\mathbf{A}(n,\sigma(n)). \tag{56}$$

The proof of the statement is by a reindexing argument which takes a bit of time to get used to.

In this formulaic definition, a typical summand is

$$\mathrm{D}(\mathbf{A})_\sigma \overset{\text{def}}{=} \varepsilon(\sigma)\cdot\mathbf{A}(1,\sigma(1))\cdots\mathbf{A}(n,\sigma(n)). \tag{57}$$

At this time, I also point out, we have by definition

$$\mathrm{D}(\mathbf{A}^\top)_{\sigma^{-1}} = \varepsilon(\sigma^{-1})\cdot\mathbf{A}^\top(1,\sigma^{-1}(1))\cdots\mathbf{A}^\top(n,\sigma^{-1}(n)). \tag{58}$$

Claim. $\mathrm{D}(\mathbf{A})_\sigma = \mathrm{D}(\mathbf{A}^\top)_{\sigma^{-1}}$

Proof of Claim. Now since σ is a permutation, and hence a bijection, $\sigma(1),\ldots,\sigma(n)$ contains each number 1 through n. Thus (57) can be rearranged, using the commutativity of R, to equal $\varepsilon(\sigma)\cdot\mathbf{A}(?_1,1)\cdots\mathbf{A}(?_1,n)$. But what are the question marks? Suppose that i_1 is the letter which gets sent to 1 by σ, i.e. suppose $\sigma(i_1)=1$. Then σ^{-1} takes 1 to i_1. Thus, in this case, the term the term $\mathbf{A}(i_1,\sigma(i_1)) = \mathbf{A}(\sigma^{-1}(1),1)$. This means that $?_1 = \sigma^{-1}(1)$. This is the case for all terms:

$$\mathrm{D}(\mathbf{A})_\sigma = \varepsilon(\sigma)\cdot\mathbf{A}(1,\sigma(1))\cdots\mathbf{A}(n,\sigma(n)) = \varepsilon(\sigma)\cdot\mathbf{A}(\sigma^{-1}(1),1)\cdots\mathbf{A}(\sigma^{-1}(n),n). \tag{59}$$

[IX] We are using our assumption that R is commutative. Notice that $\mathbf{B}(i,k)$ and $\mathbf{A}(k,j)$ are elements of R and are **not** matrices, which rarely commute.

[X] This proof is more subtle than most proofs given in this text. I would attribute most the subtleness to the awkward nature of summation notation. Moreover the result is not that important, so it's reasonable to take the result on faith. However, the method of proof becomes more important in higher level group theory courses.

We also know from Exercise 2.36 back on page 43 that $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$. Therefore (59) is

$$\varepsilon(\sigma) \cdot \mathbf{A}(\sigma^{-1}(1), 1) \cdots \mathbf{A}(\sigma^{-1}(n), n) = \varepsilon(\sigma^{-1}) \cdot \mathbf{A}(\sigma^{-1}(1), 1) \cdots \mathbf{A}(\sigma^{-1}(n), n). \tag{60}$$

Now, for our next trick. A term on the right hand side has the form

$$\mathbf{A}(\sigma^{-1}(i), i) = \mathbf{A}^\top(i, \sigma^{-1}(i))$$

So the RHS of (60) is

$$\varepsilon(\sigma^{-1}) \cdot \mathbf{A}^\top(1, \sigma^{-1}(1)) \cdots \mathbf{A}^\top(n, \sigma^{-1}(n)). \tag{61}$$

This matches exactly with (58) and we've proven our claim. ■

Since every element of $\mathcal{S}_n$ has a unique inverse the sets

$$\{\sigma : \sigma \in \mathcal{S}_n\} = \{\sigma^{-1} : \sigma \in \mathcal{S}_n\}$$

are the same. This means we have an equality of sums

$$\det(\mathbf{A}) = \sum_{\sigma \in \mathcal{S}_n} \mathrm{D}(\mathbf{A})_\sigma = \sum_{\sigma \in \mathcal{S}_n} \mathrm{D}(\mathbf{A})_{\sigma^{-1}}.$$

Moreover, by the claim: $\mathrm{D}(\mathbf{A})_{\sigma^{-1}} = \mathrm{D}(\mathbf{A}^\top)_\sigma$. Therefore, summing

$$\det(\mathbf{A}) = \sum_{\sigma \in \mathcal{S}_n} \mathrm{D}(\mathbf{A})_{\sigma^{-1}} = \sum_{\sigma \in \mathcal{S}_n} \mathrm{D}(\mathbf{A}^\top)_\sigma = \det(\mathbf{A}^\top).$$

Which is what we needed to prove. ■

Remark 3.213. It is interesting to compare **(3)** from the above with the statement for quaternions which states that

$$\overline{q_1 q_2} = \overline{q}_2 \, \overline{q}_1.$$

In fact, roughtly, the transpose in matrix theory plays a very similar role as the conjugate in quaternion theory. (Perhaps it's slightly more accurate to say that the adjoint, considered below, occupies this distinction.)

Remark 3.214. As pointed out, the proof of **(4)** is surprisingly subtle. (I wouldn't call it hard since it will, I hope, seem rather basic if you go further in group theory since you will get more used to the standard tricks of having an index of summation vary over a group). Perhaps the best way to understand it is by working a couple of concrete examples. For instance, write down the formulas for the determinant of an arbitrary three by three matrix, $\mathbf{A}$, and its transpose, $\mathbf{A}^\top$, and try to notice how you are following the proof given here as you attempt to go from one to the other.

Exercise 3.83. **T** or **F** For any two matricies **A** and **B**,

$$(\mathbf{A}\,\mathbf{B})^{\top} = \mathbf{A}^{\top}\,\mathbf{B}^{\top}.$$

Solution

Definition 3.215.

(1) A matrix **A** is *symmetric* if $\mathbf{A}^{\top} = \mathbf{A}$.

(2) A matrix **A** is *anti-symmetric* if $\mathbf{A}^{\top} = -\mathbf{A}$.

Exercise 3.84.

(1) Prove that the set of symmetric matrices in $\mathbf{M}_n(R)$ are a subspace $\mathbf{M}_n(R)$. That is prove for any two symmetric matrices **A** and **B** and scalar c then

- Prove that $\mathbf{A} + \mathbf{B}$ is symmetric.
- Prove that $c \cdot \mathbf{A}$ is symmetric.

Hint! Before doing this, look at the next part to see if perhaps you can prove both parts at the same time.

(2) Also prove that the set of anti-symmetric matrices in $\mathbf{M}_n(R)$ are a subspace $\mathbf{M}_n(R)$.

Solution

Exercise 3.85.

(1) Prove that if **A** and **B** commute and are either both symmetric or both anti-symmetric, then $\mathbf{A} \cdot \mathbf{B}$ is symmetric.

(2) Prove that if **A** and **B** commute and one is symmetric and the other is anti-symmetric, then $\mathbf{A} \cdot \mathbf{B}$ is anti-symmetric.

(3) Show by way of example that $\mathbf{A} \cdot \mathbf{B}$ needn't be symmetric if **A** and **B** are symmetric but do not commute.

(4) Prove that for any symmetric or anti-symmetric matrix **A**, that $\mathbf{A}^2$ is symmetric. Generalize to higher power of **A**.

Solution

Exercise 3.86. Prove that for any square matrix, **A** that

(1) $\mathbf{A} + \mathbf{A}^{\top}$ is symmetric.

(2) $\mathbf{A} \cdot \mathbf{A}^{\top}$ is symmetric.

Solution

(3) $\mathbf{A} - \mathbf{A}^\top$ is anti-symmetric.

For the following exercise recall Definition 2.56 for the definition of a projection.

SOLUTION

EXERCISE 3.87. Suppose R is a commutative ring where 2 is a unit.[XI] Define functions on $\mathbf{M}_n(R)$

$$\mathcal{S}(\mathbf{B}) := \frac{1}{2}\left(\mathbf{B} + \mathbf{B}^\top\right) \quad \text{and} \quad \mathcal{A}(\mathbf{B}) := \frac{1}{2}\left(\mathbf{B} - \mathbf{B}^\top\right).$$

Prove that $\mathcal{S}$ and $\mathcal{A}$ are perpendicular projections in the sense that

- $\mathcal{S}$ is a projection onto the set of symmetric matrices.
- $\mathcal{A}$ is a projection onto the set of anti-symmetric matrices.
- The map $\mathcal{A} + \mathcal{B}$ is the identity. In other words prove that for all $\mathbf{B} \in \mathbf{M}_n(R)$

$$\mathbf{B} = \mathcal{S}(\mathbf{B}) + \mathcal{A}(\mathbf{B}).$$

Hint! Using the subsequent exercise, Exercise 2.32 back on page 36 and the formulas, the only work you have to do is to prove that if $\mathbf{B}$ is symmetric then $\mathcal{S}(\mathbf{B}) = \mathbf{B}$.

We will only use the Corollary below mostly for real coefficients, but the statement and its proof work the same if we replace $\mathbb{R}$ by any commutative ring, R.

COROLLARY 3.216. *For any square matrix* $\mathbf{A}$ *(which we think of as a linear function* $R^n \to R^m$*, and any* $\mathbf{v} \in R^n$ *and* $\mathbf{w} \in R^m$ *we have that*

$$\mathbf{v} \bullet \mathbf{A}(\mathbf{w}) = \mathbf{A}^\top(\mathbf{v}) \bullet \mathbf{w}.$$

where $\bullet$ *is the usual dot product of vectors.*

PROOF. We can think of $\mathbf{v}$ and $\mathbf{w}$ are column vectors. In this case we have

$$\mathbf{v} \bullet \mathbf{w} = \mathbf{v}^\top \mathbf{w}.$$

(The right hand side is the multiplication of a 1 by n matrix and a n by 1 matrix.) Hence

$$\begin{aligned}
\mathbf{v} \bullet (\mathbf{A} \cdot \mathbf{w}) &= \mathbf{v}^\top \cdot (\mathbf{A} \cdot \mathbf{w}) \\
&= (\mathbf{v}^\top \cdot \mathbf{A}) \cdot \mathbf{w} && \text{(Associativity of matrix mult.)} \\
&= ((\mathbf{v}^\top \cdot \mathbf{A})^\top)^\top \mathbf{w} && (\top \text{ is an involution.}) \\
&= (\mathbf{A}^\top \cdot \mathbf{v})^\top \cdot \mathbf{w} \\
&= (\mathbf{A}^\top \cdot \mathbf{v}) \bullet \mathbf{w}
\end{aligned}$$

which is what we wanted to prove. ■

[XI] For our purposes, these rings are $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{Z}_n$ where n is odd.

3.5.5 Random Crap

Definition 3.217. Given an n by n matrix $\mathbf{A}$ the *trace* of $\mathbf{A}$, denoted $\mathrm{tr}\mathbf{A}$, is the sum of the diagonal entries of $\mathbf{A}$. That is

$$\mathrm{tr}\mathbf{A} = \sum_{i=1}^{n} \mathbf{A}(i,i).$$

Example 3.218. The trace of the matrix

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 2 & 1 \\ 2 & 3 & 4 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

is $1 + 2 + 4 + 1 = 8$. The trace of the matrix

$$\begin{pmatrix} [1]_8 & [2]_8 & [3]_8 \\ [1]_8 & [3]_8 & [2]_8 \\ [2]_8 & [3]_8 & [4]_8 \end{pmatrix}$$

is

$$[2]_8 + [3]_8 + [4]_4 = [9]_4 = [1]_8 .$$

Theorem 3.219. *For all* $\mathbf{A}, \mathbf{B} \in \mathbf{M}_n(R)$,

(1) *The trace is linear. That is*

- **(a)** $\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$ *and*
- **(b)** $\forall r \in R, \mathrm{tr}(r \cdot \mathbf{A}) = r \cdot \mathrm{tr}(\mathbf{A})$.

(2) $\mathrm{tr}(\mathbf{A}\,\mathbf{B}) = \mathrm{tr}(\mathbf{B}\,\mathbf{A})$.

Proof.

(1) Exercise 3.88 below.

(2)

$$\begin{aligned}
\operatorname{tr}(\mathbf{A}\cdot\mathbf{B}) &= \sum_{i=1}^{n}(\mathbf{A}\,\mathbf{B})(i,i)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{A}(i,j)\cdot\mathbf{B}(j,i)\\
&= \sum_{j=1}^{n}\sum_{i=1}^{n}\mathbf{A}(i,j)\cdot\mathbf{B}(j,i)\\
&= \sum_{j=1}^{n}\sum_{i=1}^{n}\mathbf{B}(j,i)\cdot\mathbf{A}(i,j)\\
&= \sum_{j=1}^{n}(\mathbf{B}\cdot\mathbf{A})(j,j)\\
&= \operatorname{tr}(\mathbf{B}\cdot\mathbf{A}).
\end{aligned}$$

■

WARNING! 3.220. We have seen that det preserves multiplication, that is for any two matrices **A** and **B** we have that

$$\det(\mathbf{A}\cdot\mathbf{B}) = \det(\mathbf{A})\cdot\det(\mathbf{B}),$$

and the trace preserves addition, that is for any two matrices **A** and **B** we have that

$$\operatorname{tr}(\mathbf{A}+\mathbf{B}) = \operatorname{tr}(\mathbf{A}) + \operatorname{tr}(\mathbf{B}).$$

However, neither of these matrix functions enjoy the others properties. For instance

$$\det(\mathbf{I}_2) = 1 \quad \text{and} \quad \det(2\mathbf{I}_2) = 4 \neq \det(\mathbf{I}_2) + \det(\mathbf{I}_2) = 2\det(\mathbf{I}_2)$$

which shows that the determinant is neither additive nor does it preserve scalar multiplication. Meanwhile the trace of this identity matrix is 2, so

$$2 = \operatorname{tr}(\mathbf{I}_2) = \operatorname{tr}\left(\mathbf{I}_2^2\right) \neq \operatorname{tr}(\mathbf{I}_2)\cdot\operatorname{tr}(\mathbf{I}_2) = 2\cdot 2 = 4.$$

Therefore, the trace is not multiplicative.[XII] Essentially, the only property that they do indeed have in common is that both are symmetric in the sense that neither depend on the order of their arguments, i.e. $\operatorname{tr}(\mathbf{A}\cdot\mathbf{B}) = \operatorname{tr}(\mathbf{B}\cdot\mathbf{A})$ and $\det(\mathbf{A}\cdot\mathbf{B}) = \det(\mathbf{B}\cdot\mathbf{A})$.

EXERCISE 3.88. Prove Theorem 3.219**(1)**.

SOLUTION

EXERCISE 3.89. Prove for any $\mathbf{A} \in \mathbf{M}_n(R)$ and $r \in R$ we have that

$$\det(r\mathbf{A}) = r^n \cdot \det(\mathbf{A}).$$

SOLUTION

EXERCISE 3.90. Let $\mathbf{A}$ be a 2 by 2 matrix with entries in a commutative ring. Show that we have

$$\mathbf{A}^2 = \mathrm{tr}\mathbf{A} \cdot \mathbf{A} - \det \mathbf{A} \cdot \mathbf{I}_2.$$

This is the two by two case of what is known as the *Cayley Hamilton Theorem.*

Hint! Just compute both sides of the equation and notice they are equal. Don't be intimidated by the statement which holds in many different contexts! If it makes you feel better, prove it for matrices with real coefficients and then look back at your proof and notice that you never actually used the fact that the coefficients were real!

SOLUTION

WARNING! 3.221. The formula from this previous exercise, Exercise 3.90 is only for 2 by 2 matrices! There is a more general formula which works in every dimension, but clearly our "direct check" proof given for the 2 by 2 case will not work there!

This formula is incredibly useful. For starters, using the fact that the trace is linear (Theorem 3.219**(1)**), we see that

$$\begin{aligned}
\mathrm{tr}\left(\mathbf{A}^2\right) &= \mathrm{tr}\left(\mathrm{tr}\mathbf{A} \cdot \mathbf{A} - \det \mathbf{A} \cdot \mathbf{I}_2\right) \\
&= \mathrm{tr}\left(\mathrm{tr}\mathbf{A} \cdot \mathbf{A}\right) - \mathrm{tr}\left(\det \mathbf{A} \cdot \mathbf{I}_2\right) \\
&= \mathrm{tr}\mathbf{A} \cdot \mathrm{tr}\mathbf{A} - \det \mathbf{A} \cdot \mathrm{tr}\mathbf{I}_2 \\
&= \mathrm{tr}\left(\mathbf{A}\right)^2 - \det \mathbf{A} \cdot 2 \\
&= \mathrm{tr}\left(\mathbf{A}\right)^2 - 2 \cdot \det \mathbf{A}.
\end{aligned}$$

In particular, in the case that $\det \mathbf{A} = 1$, we have that

$$\boxed{\mathrm{tr}\left(\mathbf{A}^2\right) = \mathrm{tr}\left(\mathbf{A}\right)^2 - 2.}$$

EXERCISE 3.91.

(1) Suppose that $\mathbf{A} \in \mathbf{M}_2(R)$ has trace 2 and determinant 1. Show that we have

$$\mathbf{A}^n = n\,\mathbf{A} - (n-1)\mathbf{I}_2.$$

Hint! Use induction.

SOLUTION

(2) Use the formula from the previous part to find

- $$\begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix}^{2580}$$
- $$\begin{pmatrix} [4]_5 & [1]_5 \\ [1]_5 & [3]_5 \end{pmatrix}^{1238}$$
- $$\begin{pmatrix} 1+i & -i \\ i & 1-i \end{pmatrix}^{1411}$$

(3) Let $\mathbf{A}$ be a 2 by 2 matrix with complex coefficients, determinant 1, and trace 2. Prove that if $\mathbf{A}^n = \mathbf{I}_2$ for some positive integer n, then $\mathbf{A} = \mathbf{I}_2$.

(4) Show that if a matrix $\mathbf{A}$ with real coefficients satisfies $\mathbf{A}^n = \mathbf{I}$ for some positive integer, then $\mathbf{A}$ has determinant $+1$ or -1.

(5) Lastly, suppose $\mathbf{A}$ is a 2 by 2 matrix with $\text{tr}\mathbf{A}$ not between -2 and $+2$. Prove that $\mathbf{A}^n \neq \mathbf{I}_2$ for any positive integer n.
Hint! First argue that it's enough to prove the statement if we replace $\mathbf{A}$ by $\mathbf{A}^2$ by proving that such an integer exists for $\mathbf{A}$ iff such an integer exists for $\mathbf{A}^2$. Then prove that $\text{tr}\mathbf{A}^2 > 2$ and $\det \mathbf{A}^2 = 1$. Thus, we (by replacing $\mathbf{A}$ with $\mathbf{A}^2$ if necessary) can assume that $\mathbf{A}$ has trace greater than 2 and determinant equal to 1. Next, derive the formula

$$\text{tr}\mathbf{A}^n = \text{tr}\mathbf{A} \cdot \text{tr}\mathbf{A}^{n-1} - \text{tr}\mathbf{A}^{n-2}.$$

Lastly, prove if $n_1 > n_2$ (both are integers) that $\text{tr}\mathbf{A}^{n_1} > \text{tr}\mathbf{A}^{n_2}$. Conclude that no positive power of $\mathbf{A}$ can possibly be equal to the identity.

(6) Set $\mathbf{A}$ equal to the matrix

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbf{M}_2(\mathbb{Z}_3).$$

Show that $\text{tr}\mathbf{A} = 3$ and $\mathbf{A}^4 = \mathbf{I}_2$. Why doesn't this contradict the previous part?

3.5.6 Invertible Matrices and the Adjoint

Notice the switch of i and j!

Definition 3.222. Given a square matrix $\mathbf{A}$, we define the adjoint of $\mathbf{A}$, denoted $\widetilde{\mathbf{A}}$, as the matrix with

$$\widetilde{\mathbf{A}}(i, j) = (-1)^{i+j} \det(\mathbf{A}^{(j,i)}).$$

For instance, for a 4 by 4 matrix $\mathbf{A}$

$$\begin{pmatrix} \det(\mathbf{A}^{(11)}) & -\det(\mathbf{A}^{(21)}) & \det(\mathbf{A}^{(31)}) & -\det(\mathbf{A}^{(41)}) \\ -\det(\mathbf{A}^{(12)}) & \det(\mathbf{A}^{(22)}) & -\det(\mathbf{A}^{(32)}) & \det(\mathbf{A}^{(42)}) \\ \det(\mathbf{A}^{(13)}) & -\det(\mathbf{A}^{(23)}) & \det(\mathbf{A}^{(33)}) & -\det(\mathbf{A}^{(43)}) \\ -\det(\mathbf{A}^{(14)}) & \det(\mathbf{A}^{(24)}) & -\det(\mathbf{A}^{(34)}) & \det(\mathbf{A}^{(44)}) \end{pmatrix}.$$

(Entries that get a negative sign highlighted for visual purposes.)

For a 3 by 3 matrix, let's visualize the top row:

$$\left[\det\begin{pmatrix} \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & A_{22} & A_{23} \\ \blacksquare & A_{32} & A_{33} \end{pmatrix}, -\det\begin{pmatrix} \blacksquare & A_{12} & A_{13} \\ \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & A_{32} & A_{33} \end{pmatrix}, \det\begin{pmatrix} \blacksquare & A_{12} & A_{13} \\ \blacksquare & A_{22} & A_{23} \\ \blacksquare & \blacksquare & \blacksquare \end{pmatrix}\right].$$

Notice that as we go across the row we end up going across the column. Actually in three dimensions, for a matrix with row vectors

$$\mathbf{A} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix}$$

the adjoint $\widetilde{\mathbf{A}}$ has column vectors given by the cross products[XIII] of the other two vectors, more precisely:

$$\widetilde{\mathbf{A}} = (\mathbf{r}_2 \times \mathbf{r}_3 \mid -(\mathbf{r}_1 \times \mathbf{r}_3) \mid \mathbf{r}_1 \times \mathbf{r}_2).$$

Using the formulaic version of the determinant we see that up to the choice of a sign

$$\widetilde{\mathbf{A}}(i,j) = \sum_{\sigma \in \mathcal{S}_n \sigma(j)=i} \varepsilon(\sigma) \cdot \mathbf{A}(1,\sigma(1)) \cdots \mathbf{A}(j-1,\sigma(j-1)) \cdot \mathbf{A}(j+1,\sigma(j+1)) \cdots \mathbf{A}(n,\sigma(n)).$$

The choice of the sign that we're not specifying comes about since in the sub–matrix $\mathbf{A}^{(j,i)}$ many entries get shifted about. This is rather annoying and we'll have a better way of dealing with at a later time.

Remark 3.223. Since the determinant of any matrix with coefficients in a ring R is an element of R, the adjoint of such a matrix is still a matrix with coefficients in the same ring R. In symbols

$$\mathbf{A} \in \mathbf{M}_n(R) \Rightarrow \widetilde{\mathbf{A}} \in \mathbf{M}_n(R).$$

This is **not** the case with the inverse for two reasons! First, perhaps the matrix does not have an inverse at all. Second, perhaps the inverse no longer has coefficients in R. For instance, the 1 by 1 matrix (2), $\mathbf{M}_1(\mathbb{Z})$ has inverse $(\frac{1}{2}) \ni \mathbf{M}_1(\mathbb{Z})$

Example 3.224.

(1) For any 2 by 2 matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow \widetilde{\mathbf{A}} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

[XIII] Cross products were only defined for vectors in $\mathbb{R}^3$ but the same formula can be used to define the cross product of vectors in R^3 for any ring!

(2) Let $\mathbf{A}$ denote the matrix

$$\begin{pmatrix} 2 & 1 & 3 \\ 1 & 1 & 3 \\ 2 & 1 & 1 \end{pmatrix}.$$

Then, for instance, $\mathbf{A}^{(2,3)}$ is the matrix

$$\mathbf{A}^{(2,3)} = \begin{pmatrix} 2 & 1 & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare \\ 2 & 1 & \blacksquare \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}.$$

Therefore

$$\widetilde{\mathbf{A}}(3,2) = \det \mathbf{A}^{(2,3)} = (-1)^{3+2}(2 \cdot 1 - 1 \cdot 2) = 0^{\text{XIV}}$$

(indicated in red).

To calculate another entry:

$$\mathbf{A}^{(1,2)} = \begin{pmatrix} \blacksquare & \blacksquare & \blacksquare \\ 1 & \blacksquare & 3 \\ 2 & \blacksquare & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}.$$

Therefore

$$\widetilde{\mathbf{A}}(2,1) = \det \mathbf{A}^{(1,2)} = (-1)^{2+1}(1 \cdot 1 - 3 \cdot 2) = -(-5) = 5.$$

(indicated in red).

If we do this for all entries we will see that

$$\widetilde{\mathbf{A}} = \begin{pmatrix} -2 & 2 & 0 \\ 5 & -4 & -3 \\ -1 & 0 & 1 \end{pmatrix}.$$

THEOREM 3.225. *Let* $\mathbf{A}$ *be an* n *by* n *matrix with coefficients in a ring* R*. Then*

(1) $\mathbf{A} \cdot \widetilde{\mathbf{A}} = \widetilde{\mathbf{A}} \cdot \mathbf{A} = \det(\mathbf{A}) \cdot \mathbf{I}_n$.

(2) $\mathbf{A}$ *is invertible*[XV] *iff the determinant of* $\mathbf{A}$ *is invertible in* R.

PROOF. We omit the proof of the first part. However we point out that $\mathbf{A}\widetilde{\mathbf{A}}$ upper left hand entry is, by definition, the determinant. The fact that all other entires are also the determinant is proven by using the various symmetries of the determinant. It can be proven directly in the cases ($n = 2$ and $n = 3$) in which we are most interested.

We now prove the second part. In one direction, suppose that there is an element $\det(\mathbf{A})^{-1} \in R$ so that

$$\det(\mathbf{A}) \cdot \det(\mathbf{A})^{-1} = \det(\mathbf{A})^{-1} \cdot \det(\mathbf{A}) = 1.$$

Then by this first part $\widehat{\mathbf{A}} \overset{\text{def}}{=} \det(\mathbf{A})^{-1} \cdot \widetilde{\mathbf{A}}$ satisfies Definition 3.190. Conversely, if such a matrix, $\widehat{\mathbf{A}}$, exists then

$$1 = \det \mathbf{I}_n = \det\left(\mathbf{A} \cdot \widehat{\mathbf{A}}\right) = \det \mathbf{A} \cdot \det \widehat{\mathbf{A}}$$

implies that $\det \widehat{\mathbf{A}}$ is an (multiplicative) inverse for $\det \mathbf{A}$. (In particular, one exists). ■

[XIV] Again, the swapping of indices is not a mistake!

[XV] See Definition 3.190 back on page 176

Corollary 3.226. *Suppose that* $\mathbf{A} \in \mathbf{M}_n(R)$ *is invertible. Then* $\det(\mathbf{A})^{-1} \in R$ *and*

$$\mathbf{A}^{-1} = \det(\mathbf{A})^{-1} \cdot \widetilde{\mathbf{A}}.$$

Example 3.227. If we $\mathbf{A}$ to be the matrix

$$\begin{pmatrix} 3 & 0 & 1 \\ 2 & 3 & 3 \\ 0 & 1 & 2 \end{pmatrix}$$

then we can see that the determinant of $\mathbf{A}$ is 11. Therefore $\mathbf{A}^{-1}$ makes sense for $R = \mathbb{Q}$ but not for $R = \mathbb{Z}$. Indeed a direct calculation shows that $\mathbf{A}^{-1}$ is

$$\frac{1}{11}\begin{pmatrix} 3 & 1 & -3 \\ -4 & 6 & -7 \\ -2 & -3 & 9 \end{pmatrix}.$$

If we think of the entries as elements of one of the $\mathbb{Z}_n$'s then $\mathbf{A}^{-1}$ makes sense for all n which are not divisible by 11. For instance, in $\mathbb{Z}_{13}$ we have that $[\,11^{-1}\,]_{13}$ is $[\,6\,]_{13}$. Therefore, we'd have $[\,\mathbf{A}^{-1}\,]_{13}$ is

$$[\,6\,]_{13}\begin{pmatrix} [\,3\,]_{13} & [\,1\,]_{13} & [\,10\,]_{13} \\ [\,9\,]_{13} & [\,6\,]_{13} & [\,6\,]_{13} \\ [\,11\,]_{13} & [\,10\,]_{13} & [\,9\,]_{13} \end{pmatrix} = \begin{pmatrix} [\,5\,]_{13} & [\,6\,]_{13} & [\,8\,]_{13} \\ [\,2\,]_{13} & [\,10\,]_{13} & [\,10\,]_{13} \\ [\,1\,]_{13} & [\,8\,]_{13} & [\,2\,]_{13} \end{pmatrix}.$$

Definition 3.228. The *general linear group* of degree n, denoted $\mathbf{GL}_n(R)$ is defined as the set of n by n matrices with coefficients in R whose determinant is invertible in R. We denote the set of n by n matrices whose determinant is 1 as $\mathbf{SL}_n(R)$. This latter set is called the *special linear group* of degree n.

In symbols,

$$\mathbf{SL}_n(R) = \{\mathbf{A} \in \mathbf{M}_n(R) : \det \mathbf{A} = 1\} \quad \text{and} \quad \mathbf{GL}_n(R) = \{\mathbf{A} \in \mathbf{M}_n(R) : \det \mathbf{A} \in \mathcal{U}(R)\}$$

where, as you recall, $\mathcal{U}(R)$ is the set of elements which have a (multiplicative) inverse in R. Note that $\mathbf{SL}_n(R) \subseteq \mathbf{GL}_n(R)$ since 1 is invertible in any ring (with inverse equal to itself.) For instance the matrix from the previous example is in $\mathbf{GL}_3(\mathbb{Q})$ and $\mathbf{M}_3(\mathbb{Z})$, but it is not in $\mathbf{GL}_3(\mathbb{Z})$. By the adjoint formula, we have that $\mathbf{A} \in \mathbf{SL}_n(R)$ iff $\mathbf{A}^{-1} = \widetilde{\mathbf{A}}$. Identifying 1 by 1 matrices with R itself, we have

$$\mathbf{SL}_1(R) = \{1\} \quad \text{and} \quad \mathbf{GL}_1(R) = \mathcal{U}(R).$$

(For a matrix to be in $\mathbf{GL}_n(\mathbb{Z})$ the determinant of that matrix must be ± 1 since these are the only elements of $\mathbb{Z}$ with a multiplicative inverse.) Therefore for a matrix $\mathbf{A}$ to be in, say $\mathbf{SL}_2(\mathbb{Z})$ (resp. $\mathbf{GL}_2(\mathbb{Z})$) we are requiring three things

(1) The matrix is a 2 by 2 matrix;

(2) Every entry of the matrix $\mathbf{A}$ is an integer;

(3) The determinant of the matrix is 1 (resp. ± 1).

On the other hand, for a matrix to be in $\mathbf{GL}_2(\mathbb{Q})$ we require

(1) The matrix is a 2 by 2 matrix;

(2) Every entry of the matrix $\mathbf{A}$ is a rational number;

(3) The determinant of the matrix is non-zero.

The distinction comes about because every non-zero rational number has a multiplicative inverse, but only ± 1 have multiplicative inverses over $\mathbb{Z}$. In symbols

$$\mathcal{U}(\mathbb{Z}) = \{+1, -1\} \quad \text{and} \quad \mathcal{U}(\mathbb{Q}) = \mathbb{Q}^{\times}.$$

EXAMPLE 3.229. Let's consider the matrices

(1) The zero matrix

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

is clearly not invertible for any ring.

(2) The matrix

$$\begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$$

has determinant -1 which is invertible in $\mathbb{Z}$. Since it's a 2 by 2 matrix with integer coefficients it's an element of $\mathbf{GL}_2(\mathbb{Z})$. However it is not an element of $\mathbf{SL}_2(\mathbb{Z})$ since the determinant is not 1. (Similarly it's also an element of $\mathbf{GL}_2(\mathbb{Q})$, $\mathbf{GL}_2(\mathbb{R})$, and $\mathbf{GL}_2(\mathbb{C})$.)

(3) The matrix

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

has determinant 1. Therefore it's in $\mathbf{SL}_2(\mathbb{Z})$.

(4) The matrix

$$\begin{pmatrix} 4 & 3 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

is a 3 by 3 matrix so it certainly cannot be in $\mathbf{SL}_{\textcircled{2}}$(anything) or $\mathbf{GL}_{\textcircled{2}}$(anything). It's determinant is

$$\begin{aligned} \det(4\mathbf{e}_1 + 3\mathbf{e}_2 + \mathbf{e}_3, 2\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) &= \det(4\mathbf{e}_1 + 3\mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) \\ &= \det(4\mathbf{e}_1 + 3\mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_3) \\ &= \det(4\mathbf{e}_1 + 3\mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_3) \\ &= \det(\mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_3) \\ &= \det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \\ &= 1. \end{aligned}$$

Thus it is an element of $\mathbf{SL}_3(\mathbb{Z})$ (and $\mathbf{GL}_3(\mathbb{Z})$, $\mathbf{SL}_3(\mathbb{Q})$, $\mathbf{GL}_3(\mathbb{Q})$ etc.).

(5) The matrix

$$\begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

does not have integer coefficients so it can't possibly be in $\mathbf{SL}_2(\mathbb{Z})$. It does however have determinant 1 and rational coefficients, so it's an element of $\mathbf{SL}_2(\mathbb{Q})$, $\mathbf{SL}_2(\mathbb{Z})$, and $\mathbf{SL}_2(\mathbb{C})$ as well as the $\mathbf{GL}_2()$ versions.

(6) The matrix

$$\begin{pmatrix} 2 & 2 \\ 1 & 6 \end{pmatrix}$$

has determinant 10 and is therefore not in $\mathbf{SL}_2(\mathbb{Z})$ or $\mathbf{GL}_2(\mathbb{Z})$. It is however in $\mathbf{GL}_2(\mathbb{Q})$ etc. It descends to an element of $\mathbf{SL}_2(\mathbb{Z}_3)$ and $\mathbf{SL}_2(\mathbb{Z}_5)$, but does not descend to an element of $\mathbf{GL}_2(\mathbb{Z}_2)$ or $\mathbf{GL}_2(\mathbb{Z}_5)$. It descends to an element of $\mathbf{GL}_2(\mathbb{Z}_7)$ but not an element of $\mathbf{SL}_2(\mathbb{Z}_7)$.

3.5.7 Closure Properties

The following theorem should be read: $\mathbf{SL}_n(R)$ is closed under multiplication and inversion.

Theorem 3.230. *Let n be any positive integer and R be any commutative ring. Then if $\mathbf{A}, \mathbf{B} \in \mathbf{SL}_n(R)$, then*

(1) $\mathbf{A} \cdot \mathbf{B} \in \mathbf{SL}_n(R)$

(2) $\mathbf{A}^{-1} \in \mathbf{SL}_n(R)$.

Proof.

(1) Let $\mathbf{A}, \mathbf{B} \in \mathbf{SL}_n(R)$. Then, by definition, $\det \mathbf{A} = 1$ and $\det \mathbf{B} = 1$. Therefore

$$\begin{aligned} \det(\mathbf{A} \cdot \mathbf{B}) &= \det \mathbf{A} \cdot \det \mathbf{B} && \text{(Theorem 3.205)} \\ &= 1 \cdot 1 && (\det \mathbf{A} = 1, \det \mathbf{B} = 1) \\ &= 1. \end{aligned}$$

Therefore $\mathbf{A} \cdot \mathbf{B} \in \mathbf{SL}_n(R)$ by definition and we are done.

(2) By Corollary 3.208 we know that $\det\left(\mathbf{A}^{-1}\right) = \det(\mathbf{A})^{-1}$. But, under the assumption that $\det \mathbf{A} = 1$, the right hand side of this equation is 1^{-1} which is 1. Therefore $\det\left(\mathbf{A}^{-1}\right) = 1$ has determinant 1. We still haven't proven the subtle part of this this, namely that the coefficients of $\mathbf{A}^{-1}$ still live in the same ring, R (c.f. Example 3.227). But for any matrix, we have that

$$\mathbf{A} \cdot \widetilde{\mathbf{A}} = \widetilde{\mathbf{A}} \cdot \mathbf{A} = \det \mathbf{A}$$

where $\widetilde{\mathbf{A}}$ is the adjoint of $\mathbf{A}$. Since the adjoint of any matrix still has coefficients in the same ring, see Remark 3.223, and since $\det \mathbf{A} = 1$, we see that $\widetilde{\mathbf{A}} = \mathbf{A}^{-1} \in \mathbf{SL}_n(R)$.

■

It's just as easy to show that $\mathbf{GL}_n(R)$ also has these properties, but we do this in a much more general context later. For now we record a different theorem about $\mathbf{GL}_n(R)$. In words, it says that a matrix is invertible (in $\mathbf{M}_n(R)$) iff its determinant is invertible (in R).

Theorem 3.231.

$$\mathcal{U}(\mathbf{M}_n(R)) = \mathbf{GL}_n(R).$$

Proof. This is simply a restatement of Theorem 3.225(**2**). ∎

In words this says

A matrix is invertible over R iff its determinant is invertible in R.

Example 3.232. The matrix

$$\mathbf{A} := \begin{pmatrix} 3 & 7 \\ 7 & 17 \end{pmatrix}$$

which corresponds to the linear transformation

$$\mathbf{A}^\flat : (x, y) \mapsto (3x + 7y, 7x + 17y)$$

has determinant $6 - 4 = 2$. Therefore, since the number 2 is invertible over $\mathbb{Q}$, the matrix $\mathbf{A}$ is invertible over $\mathbb{Q}$, i.e. the map

$$\begin{aligned} \mathbf{A}^\flat &: \mathbb{Q}^2 \to \mathbb{Q}^2 \\ &: (x, y) \mapsto (3x + 7y, 7x + 17y) \end{aligned}$$

is a bijection. However, since the number 2 is not invertible over $\mathbb{Z}$ since $1/2$ is not an integer, the matrix $\mathbf{A}$ and the corresponding linear transformation

$$\begin{aligned} \mathbf{A}^\flat &: \mathbb{Z}^2 \to \mathbb{Z}^2 \\ &: (x, y) \mapsto (3x + 7y, 7x + 17y) \end{aligned}$$

are **not** a bijection. Since the latter is a restriction of a former, like the restriction of any injective function, the map $\mathbf{A}^\flat : \mathbb{Z}^2 \to \mathbb{Z}^2$ is injective, therefore since it's not a bijection, it cannot possible be surjective. Let's verify this explicitly.

Is $\mathbf{e}_1 \in \operatorname{Im}(\mathbf{A}^\flat)$ when we think of it a map between ordered pairs of integers? In other words, does there exist integers x and y which solve the equations

$$\begin{aligned} 1 &= 3x + 7y \\ 0 &= 7x + 17y. \end{aligned}$$

Well, we can subtract twice the first equation from the second to see that

$$\begin{aligned} 1 &= 3x + 7y \\ -2 &= x + 3y. \end{aligned}$$

Then we can subtract 3 times the new second from the first to see that

$$\begin{aligned} 7 &= \quad - 2y \\ -2 &= x + 3y. \end{aligned}$$

This implies that y must equal $-7/2$, which is not an integer. This proves that $\mathbf{e}_1$ is not in the image, verifying what we already knew: the map is not surjective.

3.5.8 Some Important Matrices

Of course the identity matrix, $\mathbf{I}$ and the zero matrix, $\mathbf{0}$, are important matrices. Another important matrices are the so called elementary matrices.

Definition 3.233. Let i and j be integers. Define the matrix $\mathbf{E}_{ij}$ as the matrix whose only non–zero entry is a 1 in the i^{th} row and j^{th} column.

Exercise 3.92. Prove that

$$\mathbf{E}_{ij} \cdot \mathbf{E}_{kl} = \delta_{jk}\mathbf{E}_{il}$$

where δ is the Kroneker delta, i.e. takes the value 1 iff $i = j$ and takes the value 0 otherwise.

SOLUTION

Remark 3.234. As usual, we are being sloppy. These matrices depend on the ring, and the degree in question. For instance, over $\mathbb{Z}$ and degree 4 the matrix

$$\mathbf{E}_{41} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Definition 3.235. If $i < j$ define $\mathbf{L}_{ij}$ as $\mathbf{I} + \mathbf{E}_{ij}$ and $\mathbf{R}_{ij}$ as $\mathbf{L}_{ij}^{\top}$ (which equals $\mathbf{I} + \mathbf{E}_{ji}$). Given distinct integers, Define $\mathbf{LR}_{ij}$ to be $\mathbf{L}_{ij}$ if $i < j$ and $\mathbf{R}_{ji}$ if $j < i$.

For instance,

$$\mathbf{R}_{41} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Exercise 3.93. Prove that $\mathbf{LR}_{ij}$ and $\mathbf{LR}_{kl}$ commute iff

$$(j \neq k \quad \text{and} \quad i \neq l) \quad \text{or} \quad (i = k \quad \text{and} \quad j = l).$$

Hint! Prove that $\mathbf{LR}_{ij}$ and $\mathbf{LR}_{kl}$ commute iff $\mathbf{E}_{ij}$ and $\mathbf{E}_{kl}$ commute.

SOLUTION

Exercise 3.94. Prove that for any integer n, $\mathbf{LR}_{ij}^{n} = \mathbf{I} + n\mathbf{E}_{ij}$.

SOLUTION

Exercise 3.95. For matrix thought of as a collection of rows,

$$\mathbf{A} = \begin{pmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_n \end{pmatrix}$$

prove that

$$\mathbf{L}_{ij}\mathbf{A} = \begin{pmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_{i-1} \\ \mathbf{r}_i + \mathbf{r}_j \\ \mathbf{r}_{i+1} \\ \vdots \\ \mathbf{r}_j \\ \vdots \\ \mathbf{r}_n \end{pmatrix} \quad \text{and} \quad \mathbf{R}_{ij}\mathbf{A} = \begin{pmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_i \\ \vdots \\ \mathbf{r}_{j-1} \\ \mathbf{r}_j + \mathbf{r}_i \\ \mathbf{r}_{j+1} \\ \vdots \\ \mathbf{r}_n \end{pmatrix}.$$

and if instead we think of $\mathbf{A}$ as a column matrix

$$\begin{pmatrix} \mathbf{c}_1 | & \cdots & | \mathbf{c}_n \end{pmatrix}$$

then

$$\mathbf{A} \cdot \mathbf{L}_{ij} = \begin{pmatrix} \mathbf{c}_1 | & \cdots & | \mathbf{c}_i | & \cdots & | \mathbf{c}_j + \mathbf{c}_i | & \cdots & | \mathbf{c}_n \end{pmatrix}, \text{ and}$$
$$\mathbf{A} \cdot \mathbf{R}_{ij} = \begin{pmatrix} \mathbf{c}_1 | & \cdots & | \mathbf{c}_i + \mathbf{c}_j | & \cdots & | \mathbf{c}_j | & \cdots & | \mathbf{c}_n \end{pmatrix}.$$

Hint! To save yourself a bit of time, note that

$$\left(\begin{pmatrix} \mathbf{c}_1 | & \cdots & | \mathbf{c}_n \end{pmatrix} \cdot \mathbf{R}_{ij}. \right)^\top = \mathbf{L}_{ij} \cdot \begin{pmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_n \end{pmatrix} \quad \text{and} \quad \left(\begin{pmatrix} \mathbf{c}_1 | & \cdots & | \mathbf{c}_n \end{pmatrix} \cdot \mathbf{L}_{ij}. \right)^\top = \mathbf{R}_{ij} \cdot \begin{pmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_n \end{pmatrix}.$$

In the case of 2 by 2 matrices, we will typically set

$$\mathfrak{L} = \mathbf{L}_{12} \quad \text{and} \quad \mathfrak{R} = \mathbf{R}_{12}.$$

Definition 3.236. A *scalar matrix* in $\mathbf{M}_n(R)$ is a matrix of the form $r\mathbf{I}_n$ for some $r \in R$. A matrix $\mathbf{A}$ is a *diagonal matrix* if

$$i \neq j \Rightarrow \mathbf{A}(i, j) = 0.$$

Aptly named, a diagonal matrix is a matrix whose only non-zero entries lie along the diagonal. Of course a scalar matrix is a diagonal matrix. The converse is plainly false. For instance

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

is diagonal but not scalar.

DEFINITION 3.237.

(1) For $\mathbf{r} = (r_1, \ldots, r_n) \in R^n$, define the n by n matrix, denoted

$$\delta(\mathbf{r}),$$

as the diagonal matrix with $(r_1, \ldots, r_n)$ along its diagonal.

(2) Define $\Delta(n, R)$ as the set of all diagonal matrices and $\mathcal{S}(n, R)$ as the subset of scalar matrices in $\mathbf{M}_n(R)$.

In our notation, the scalar matrix

$$r\mathbf{I}_n = \delta((r, \ldots, r)).$$

Scalar matrices are easy to understand in terms of linear maps. They correspond to scalar multiplication by r:

$$r\mathbf{I}_n : \mathbf{v} \mapsto r \cdot \mathbf{v}.$$

DEFINITION 3.238. For a subset of $G \subseteq \mathbf{M}_n(R)$ define ΔG (resp. $\mathcal{S}G$) as the set of diagonal (resp. scalar) matrices in G, that is $\Delta G \overset{\text{def}}{=} G \cap \Delta(n, R)$ (resp. $\mathcal{S}G \overset{\text{def}}{=} G \cap \mathcal{S}(n, R)$).

Of course we have $\mathcal{S}G \subseteq \Delta G \subseteq G$. Important special cases include $\Delta\mathbf{GL}_n(R)$, $\Delta\mathbf{SL}_n(R)$, $\mathcal{S}\mathbf{GL}_n(R)$, and $\mathcal{S}\mathbf{SL}_n(R)$. Note that

$$\Delta\mathbf{SL}_2(R) = \left\{ \begin{pmatrix} r & 0 \\ 0 & \frac{1}{r} \end{pmatrix} : r \in \mathcal{U}(R) \right\}.$$

EXAMPLE 3.239. For instance

$$\delta((1, -2, 5)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

Of course $\delta((r_1, \ldots, r_n))$ is invertible iff all of the r_i are invertible, in which case the inverse is

$$\delta((r_1^{-1}, r_2^{-1}, \ldots, r_n^{-1})).$$

Moreover

$$\delta((r_1, r_2 \ldots, r_n)) \cdot \delta((s_1, s_2 \ldots, s_n)) = \delta((r_1 \cdot s_1, r_2 \cdot s_2, \ldots, r_n \cdot s_n)) \quad \text{and} \quad r\mathbf{I}_n \cdot s\mathbf{I}_n = rs\mathbf{I}_n = s\mathbf{I}_n \cdot r\mathbf{I}_n.$$

We also have

$$\delta((r_1,\dots,r_n))\cdot\begin{pmatrix}\mathbf{r}_1\\ \vdots\\ \mathbf{r}_n\end{pmatrix}=\begin{pmatrix}r_1\mathbf{r}_1\\ \vdots\\ r_n\mathbf{r}_n\end{pmatrix}\quad\text{and}\quad(\mathbf{c}_1\mid\cdots\mid\mathbf{c}_n)\cdot\delta((r_1,\dots,r_n))=(r_1\mathbf{c}_1\mid\cdots\mid r_n\mathbf{c}_n),$$

and if $\delta((r_1,\dots,r_n))$ is invertible

$$\delta((r_1,\dots,r_n))\cdot\begin{pmatrix}A_{11}&A_{12}&\cdots&A_{1n-1}&A_{1n}\\ A_{21}&A_{22}&\cdots&A_{2n-1}&A_{2n}\\ \vdots&\vdots&\vdots&&\\ A_{n-11}&A_{n-12}&\cdots&A_{n-1n-1}&A_{n-1n}\\ A_{n1}&A_{n2}&\cdots&A_{nn-1}&A_{nn}\end{pmatrix}\cdot\delta((r_1,\dots,r_n))^{-1}=\begin{pmatrix}A_{11}&\frac{r_1}{r_2}A_{12}&\cdots&\frac{r_1}{r_{n-1}}A_{1n-1}&\frac{r_1}{r_n}A_{1n}\\ \frac{r_2}{r_1}A_{21}&A_{22}&\cdots&\frac{r_2}{r_{n-1}}A_{2n-1}&\frac{r_2}{r_n}A_{2n}\\ \vdots&\vdots&\vdots&&\\ \frac{r_{n-1}}{r_1}A_{n-11}&\frac{r_{n-1}}{r_2}A_{n-12}&\cdots&\frac{r_{n-1}}{r_{n-1}}A_{n-1n-1}&\frac{r_{n-1}}{r_n}A_{n-1n}\\ \frac{r_n}{r_1}A_{n1}&\frac{r_n}{r_2}A_{n2}&\cdots&\frac{r_n}{r_{n-1}}A_{nn-1}&A_{nn}\end{pmatrix}.$$

A clearner way of saying this is via

$$\left(\delta(\mathbf{r})\cdot\mathbf{A}\cdot\delta(\mathbf{r})^{-1}\right)(i,j)=\frac{r_i}{r_j}\mathbf{A}(i,j).$$

4 Algebraic Systems

At this point, we can now add the word "abstract" to our algebra. We no longer study things on a case-by-case basis but we study them by seeing exactly what their various properties imply.

Essentially, *algebra* is the study of sets and binary operations. Together they form what is known as an algebraic system. The formal definition is

Definition 4.1. Let M be a set.

(1) A function

$$\star : M \times M \to M$$

is called a *binary operation* (or just an operation) or a *multiplication* defined on M.

(2) An *algebraic system* is a pair $(M, \star)$ where M is a set and $\star$ is a binary operation defined on M.

Remark 4.2. People with a background in computer science like to call algebraic system a *magma*. For instance, this is the case in `GAP`.

4.1 First Examples

For example, if $M = \mathbb{Z}$, then $\star = \cdot$, the usual multiplication of integers, defines a binary operator on $\mathbb{Z}$ and $(\mathbb{Z}, \cdot)$ is an algebraic system. In this situation, we have $\cdot(11, 7) = 77$. Of course we do not write binary operations in this way.

Notation 4.3. For a binary operation $\star : M \times M \to M$ we write $m_1 \star m_2$ instead of $\star(m_1, m_2)$. Generically, we read this as "m_1 times m_2."

As pointed out before giving the notation, this notation of course matches what we usually do. For instance, we write $11 \cdot 7 = 77$ instead of $\cdot(11, 7) = 77$ and $11 + 5 = 16$ instead of $+(11, 5) = 16$.

WARNING! 4.4. Several things which are commonly called *multiplications* or *products* in mathematics do not fit this description. For instance, the dot product of two vectors in $\mathbb{R}^3$ is a function $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, i.e. it takes to vectors and returns a *number* so the output type is not the same as the input types and so this does not qualify as a binary operation.

Scalar multiplication

$$c \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} cv_1 \\ cv_2 \\ cv_3 \end{bmatrix}$$

is also not a binary operation since the two input types (a real number and a three vector) are not of the same type.

The following definition should be expected. It's what we've talked about quite a lot already.

DEFINITION 4.5. Given an algebraic system $(M, \star)$ and a subset $N \subseteq M$ we say that N *is closed under* $\star$ if

$$n, \widehat{n} \in N \Rightarrow n \star \widehat{n} \in N.$$

Written out, a binary operation is a rule that assigns every (ordered) pair of elements in a set M (i.e. a single element of $M \times M$) a unique element of that same set M. Let's highlight the hidden nature of this definition.

WARNING! 4.6. Every time I teach this course I see students write something along the lines of $M \star M$ or $M \star M \to M$ at some point. I have no idea why they do this, but it's complete and utter non-sense! The set $M \times M$ is by definition the set of ordered pairs, $\{(m_1, m_2) \in m_1, m_2 \in M\}$. The $\times$ which appears there has **nothing, NOTHING, NOTHING!** to do with any sort of multiplication.

(1) Note that by definition, if $(M, \star)$ is an algebraic system, then M is closed under $\star$.

(2) A binary operation must be defined everywhere. For instance, $(\mathbb{R}, \div)$ is not an algebraic system since the ordered pair $(1, 0)$ is not in the domain of $\div$. This is just a fancy way of saying that $1/0$ is not defined.

(3) Since functions take "equals to equals" we can always do "the same thing to both sides of any equation." Thus if we have an equation $y = z$ and $x \in M$, then we must have that $x \star y = x \star z$. This is because in this case the ordered pair $(x, y) = (x, z)$ and so using the single valued nature, which is present since $\star$ is by definition a function,

$$x \star y = \star(x, y) = \star(x, z) = x \star z.$$

(4) In general, we will use the JFK rule when dealing with binary operations, we typically do not ask why something *is* a binary operation, we only ask ourselves why it *is not*. If we cannot come up with a good reason for that then we just kinda go with it. This is akin to proving that a particular relation, like $x \mapsto x^4 + x + 1$ is single valued: it almost seems "too obvious to prove."

EXAMPLE 4.7.

(1) As usual, the third item above is typically too trivial to even prove! It often holds by the "why wouldn't it be" reasoning we often use to prove that a written rule is single valued.

Therefore, in practice, to prove that something is a binary operation we only really need to check the first two things. Besides these three trivial facts–which are just restatements that a binary operation is a function–there is essentially nothing else that is true for an abstract binary operation. For instance, we are not assuming a binary operator is associative or commutative. For instance, something as crazy looking as

$$x \star y := \frac{\left(x^5 + xy + \sin x + y^{5280}\right)^6}{x^2 + y^2 + 1}$$

is a binary operation on the set $\mathbb{R}$.

However, it is not a binary operation on the set $\mathbb{Q}$ since

$$1 \star 0 = \frac{1}{2}\left(1 + \sin(1)\right)^6 \notin \mathbb{Q}$$

and so $\mathbb{Q}$ is not closed under $\star$. It is also not a binary operation on $\mathbb{C}$ since $0 \star i$ (as well as $i \star 0$) are not defined.

(2) In our current language, a big chunk of what we have been doing in our current language has been to show that certain sets are algebraic systems under a given binary operation.

System	$(\mathbb{S}^3, *)$	$(\mathbb{S}^1, *)$	$(\mathbf{SL}_n(\mathbb{C}), \cdot)$	$(\mathcal{S}_n, \circ)$	$(\mathcal{F}_n, \circ)$
Thm	3.101(1)	3.37(3)	3.230(1)	2.57(1)	2.48(2)

.

We will greatly expand this table even in this tiny chapter.

REMARK 4.8. As the text moves on we will slowly incorporate the following convention:

CONVENTION 4.9. Whenever possible, we refer to an algebraic system $(M, \star)$ simply as the set M itself.

Initially, this convention may cause a lot of trouble, I will do my best to avoid it! For instance the statement, "the algebraic system $\mathbb{Z}$" may cause a bit of confusion since it's not clear at all which operation I am referring to: addition or multiplication. Some sets, like the set of function $\mathbb{R} \to \mathbb{R}$ have three reasonable binary operations that we could consider: addition of functions, multiplication of functions, or composition of functions!

EXERCISE 4.1. Which of the following are algebraic systems?

SOLUTION

(1) $(\mathbb{N}, +)$
(2) $(\mathbb{N}, -)$
(3) $(\mathbb{Z}, +)$
(4) $(\mathbb{Z}, \div)$
(5) $(\mathbb{Z}^\times, \div)$
(6) $(\mathbb{Q}, \div)$
(7) $(\mathbb{Q}^\times, \div)$
(8) $(\mathcal{P}(X) \setminus \{\varnothing\}, \cup)$
(9) $(\mathcal{P}(X), \setminus)$
(10) $(\mathcal{P}(X) \setminus \{\varnothing\}, \setminus)$
(11) $(\{\{5\}\}, \cdot)$
(12) $(\{\{5\}\}, +)$
(13) $(\mathcal{F}_X, \circ)$
(14) $(\{0, 1\}, \cdot)$
(15) $(\{0, 1\}, +)$
(16) $(\mathbb{S}^3, *)$
(17) $(\mathbb{R}^3, \times)$
(18) $(\mathbb{R}^3, \bullet)$
(19) $(\mathbb{Z}_5, \cdot)$
(20) $(\mathbb{Z}_6, \cdot)$
(21) $(\mathcal{F}\mathrm{un}(X, Y), \circ)$

Here, X is a non-empty set. Recall $\mathcal{P}(X)$ is the power set of X, $\setminus$ denotes the set difference of two sets, and $\{\{n\}\}$ denotes the n^{th} roots of unity.

EXERCISE 4.2. Define $\oplus$ via

$$\begin{aligned} \oplus &: \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q} \\ &: \frac{a_1}{a_2} \oplus \frac{b_1}{b_2} \mapsto \frac{a_1 + b_1}{a_2 + b_2}. \end{aligned}$$

(1) Is $\oplus$ a binary operation?

(2) Is $(\mathbb{Q}, \oplus)$ an algebraic system?

SOLUTION

For small enough sets M this can be seen as a multiplication table which is obtained by replacing the (m_1, m_2) entry of the table for $M \times M$ with $m_1 \star m_2$.

TABLES READ COLUMN ENTRY ⋆ ROW ENTRY

EXAMPLE 4.10. If M contains only a single element, say $M = \{e\}$, then the only possible algebraic system on M is

$\star$	e
e	e

.

In this algebraic system we have $e \star e = e$. We typically denote this algebraic system simply as $\mathfrak{e}$[I].

EXAMPLE 4.11. Define

$$M = \{e, a, b\}$$

[I] We use this notation for any algebraic system with a single element.

and $\star$ according to the following table

$\circ$	e	a	b
e	e	a	b
a	a	e	e
b	b	e	b

then we have, for instance, $a \star b = e$ and $b \star e = b$.

Example 4.12. If on the other hand, we defined $M = \{e, a, b, c\}$ and $\star$ instead according to this table

$\star$	e	a	b	c
e	c	a	b	c
a	a	e	e	a
b	e	e	b	e
c	c	a	a	b

then we have that $b \star c = e$ and $c \star b = a$.

To see a more complicated calculation we have

$$\begin{aligned}((a \star c) \star (c \star b)) \star a &= (a \star a) \star a \\ &= e \star a \\ &= a.\end{aligned}$$

Exercise 4.3. Consider the algebraic system

$\star$	e	a	b	c
e	a	e	a	c
a	e	e	c	a
b	c	b	b	e
c	a	b	e	b

Calculate

(1) $(a \star b) \star (c \star e)$

(2) $((a \star b) \star c) \star e$

(3) $a \star ((b \star c) \star e)$

(4) $(a \star (b \star c)) \star e$.

Hint! All four are different!

SOLUTION

General algebraic systems simply have too little structure to even bother making a serious study of them. Essentially, nothing is true in general about them. This is because we could just randomly fill in the table with elements of M without any regard what so ever! So, if M were to contain n elements, there would be a total of $n \cdot n^2 = n^3$ possible algebraic systems which can be defined on M (each of the n^2 entries of the table can be filled in a total of n different ways.) One condition we can impose on an algebraic system is that it be associative. This means what you would expect it to mean:

4.2 Types of Algebraic Systems: Semi-Groups, Monoids, Loops, and Groups!

Definition 4.13. A binary operator $\star$ defined on a set M is said to be associative if for all a, b and $c \in M$ we have that

$$(a \star b) \star c = a \star (b \star c).$$

In which case we say that $(M, \star)$ is an *associative algebraic system* or (less descriptively) a *semi-group*.

Remark 4.14. Negating what it means for a binary operation to be non-associative requires that we simply find some a, b, and c so that

$$(a \star b) \star c \neq a \star (b \star c).$$

Warning! 4.15. I didn't say that a, b, and c were distinct. For instance, in the algebraic system

$\star$	a	b
a	b	a
b	b	a

we have

$$(a \star a) \star a = b \star a = b \neq a = a \star b = a \star (a \star a).$$

Exercise 4.4. Is the following algebraic system associative?

Solution

$\star$	a	b
a	b	a
b	a	a

Notice that showing that an algebraic system is associative is a pain. Even in the simple simple case that the set in question has only two elements $\{a, b\}$. There we would have to verify that

$$\begin{aligned} a \star (a \star a) &= (a \star a) \star a & a \star (a \star b) &= (a \star a) \star b \\ a \star (b \star a) &= (a \star b) \star a & a \star (b \star b) &= (a \star b) \star b \\ b \star (a \star a) &= (b \star a) \star a & b \star (a \star b) &= (b \star a) \star b \\ b \star (b \star a) &= (b \star a) \star a & b \star (b \star b) &= (b \star b) \star b \end{aligned}$$

a total of 8 steps (either 8 or 16 calculations, depending on how you look at it).

The following two exercises show that picking algebraic systems at random is a terrible way to find associative ones. The second problem shows that the probability drops off as the number of elements increase incredibly fast. The first problem shows that even if we are lucky enough to randomly pick an associative algebraic system, the number of calculations we would need to make would quickly become unnerving.

EXERCISE 4.5. In general, to show that an algebraic system with (whose associated set has) 3 elements is associative, how many equalities do you need to verify? How many when it has n elements?

SOLUTION

MAPLE EXERCISE 4.6. For a fixed positive integer n we write a code in MAPLE which allows us to guess the probability that a random algebraic system with n elements be associative.
Hint! It's easiest to think of algebraic system as an n by n matrix whose entries are forced to take on the values in $\overline{n}$.

SOLUTION

```
> with(LinearAlgebra):
> ASSC:=proc(M,n)
    local flag, i, j, k;
    flag:=0;
    for i from 1 to n do for j from 1 to n do for k from 1 to n do
    flag:=flag+abs(M[M[i,j],k]-M[i,M[j,k]]);
    od; od; od;
    return(flag);
  end;
>CT:=proc(m,n)
    local i, j, M, count;
    count:=0;
    for i from 1 to m do
    M:=RandomMatrix(n,n,generator=rand(1..n));
    if(ASSC(M,n)=0) then count:=count+1;
    fi;
    od;
    return(count);
  end;
```

The first command loads the LinearAlgebra package which is necessary to load in the `RandomMatrix` command. The process which appears after this, `ASSC` tests how many non-associative triples there are in a given algebraic system M (i.e. 3 by 3 matrix). In particular M is associative iff `ASSC(M)` is zero. Next to get a better grasp on probabilities we will need to use the command `RandomMatrix`. The second process `CT(m,n)` then counts how many randomly selected n by n matrices correspond to associative algebraic systems out of a total of m. For me in the case that $n = 2$ there were a total of $5,036$ out of $10,000$ which *were* associative (it's not difficult to see that there are a total of 8 associative algebraic systems on two elements!), for $n = 3$ there were a total of 52 out of $10,000$ which *were* associative and for $n = 4$ I didn't find **ANY** after running through $200,000$ random algebraic systems!

I hope that these last two problems convince you that picking algebraic systems at random and studying them by their tables alone blows. In practice, however, 95% of the time we will prove that an algebraic system is contained inside of a larger algebraic system which is associative. For instance, after going through the trouble of proving that $(\mathbb{H}, *)$ is associative, there was no need to prove that $(\mathbb{S}^3, *)$ was also associative. Another example that comes up quite often is a set of matrices. We will never actually need to show that these systems are associative, we simply need to use the fact that matrix multiplication is associative in general.

Let's investigate this avenue further. First, how do we make for an algebraic system to be contained in another: we want more than just a subset, we want that the operations be the same. Therefore, for instance, the non-zero real numbers under multiplication are not a sub–algebraic system of the real numbers under addition. However, the non-zero rational numbers under multiplication are a sub algebraic system of the real numbers under multiplication. The key condition which is hidden here is that the operation restrict to an operation. This means if $A \subseteq X$ and X has an operation $\star$ that the restriction of $\star$ to A is also a binary operation. This means for all a_1 and a_2 in A that there exists a unique element $a_1 \star a_2$ which **lives in A**. As you might have guessed by the obnoxious use of bold text the only non-trivial thing here is that it **lives in** A. We refer to this property as *closure* .

Definition 4.16. Let $A \subseteq M$ where M is an algebraic system with binary operation $\star$. Then A is a *sub-algebraic system* if it is closed under $\star$.

Example 4.17. Consider the algebraic system $(\mathbb{Z}, +)$. Then the even numbers are a subalgebraic system since the sum of two even numbers is even. However, the odd numbers are not a subalgebraic system. For instance, $1 + 1 = 2$ is even (hence not odd) but 1 is odd. On the other hand, if we consider the algebraic system $(\mathbb{Z}, \cdot)$, then both the even numbers and the odd numbers are algebraic systems. This is simply because the product of two even (odd) numbers is even (odd).

You should think of sub-algebraic systems as playing the same role as subsets play when studying general sets.

Theorem 4.18. *A sub-algebraic system of a semi-group is a semi-group.*

Proof. Obvious. ∎

The reason the proof is so obvious is because the quantifiers involved in associativity are "for all's" and properties involving them are always inherited by subsets. For instance, if every triple of people at UIC get along, then obviously every triple of people taking any particular course at UIC also get along. On the other hand, one needs to be more careful existence quantifiers. if we only knew that there existed a genius at UIC we would not know that there would be a genius in a particular classroom.

Example 4.19. The usual additions and multiplications are associative. The enemy of every high school student is that subtraction and division are not. It actually matters where you put the parenthesis with these operations! This is represented by the fact that you need to distribute the minus sign! In other words, in general,

$$a - (b - c) \neq (a - b) - c.$$

For instance, $1-(1-1)=1\neq -1=(1-1)-1$. The fact that divison is not associative is represented by the fact that in general

$$\frac{a}{b/c}\neq\frac{a/b}{c}.$$

For instance, $(1/2)/3=1/6$ but $1/(2/3)$ is $3/2$. The cross product in $\mathbb{R}^3$, $\times$, is also not associative. For instance,

$$(\mathbf{i}\times\mathbf{i})\times\mathbf{j}=\mathbf{0}\neq-\mathbf{j}=\mathbf{i}\times\mathbf{k}=\mathbf{i}\times(\mathbf{i}\times\mathbf{j}).$$

We have shown that the cross product can, in a sense, be corrected by adding a dimension to give us $(\mathbb{H},*)$ which *is* associative.

Exercise 4.7. (First Name Operation) Let M be any set and define an algebraic system on M by defining

$$a\star b=a.$$

(That is every product of two elements of M is defined to be the first element of M). Show that $(M,\star)$ is an associative algebraic system (i.e. a semi-group).

SOLUTION

Exercise 4.8. (Constant Algebraic System) Let M be any non-empty set containing some element m_0. Then define an algebraic system $(M,\star)$ via

$$a\star b=m_0.$$

Show that $(M,\star)$ is an associative algebraic system (i.e. a semi-group).

SOLUTION

Exercise 4.9. Prove the algebraic systems in Example 4.26 are associative.

SOLUTION

Associativity says that a product $a\star b\star c$ is unambiguous.[I] One can prove by using an annoying induction (we will prove this in a detour while proving something more abstract) that this statement generalizes to higher product. Thus each of the five ways ways of associated four elements

$$a\star((b\star c)\star d)$$
$$(a\star(b\star c))\star d$$
$$(a\star b)\star(c\star d)$$
$$((a\star b)\star c)\star d$$
$$a\star(b\star(c\star d))$$

[I] A fancy way of saying this is *the function*

$$:M\times M\times M\to M$$
$$:(a,b,c)\mapsto a\star b\star c$$

is well defined.

IN A SEMI-GROUP NO NEED FOR PARENTHESIS.

give the same result. This is proof by annoyance. This is shown as follows

$$a \star (b \star (c \star d)) = a \star ((b \star c) \star d) = (a \star (b \star c)) \star d = ((a \star b) \star c) \star d = (a \star b) \star (c \star d).$$

Notice at every stage I only re-associated a triple. For instance, in the last step the triple consisted of $a \star b$, c and d. The general proof for products of more than $n > 4$ is just as annoying, but it is not too much more difficult. (The most difficult aspect is establishing exactly what it is you're trying to prove and the notation to prove it!) You can either do this as an exercise, look in a separate book, or check out Theorem B.2 in the appendix for a more abstract two birds with one stone proof. We record this as a theorem for now.

THEOREM 4.20 (Generalized Associativity). *In an associative algebraic system you can re-associate any product of a finite string of elements.* ■

Therefore from now on, when we are in an associative situation we can unambiguously write any product of n-terms

$$a_1 \star \cdots \star a_n$$

without parenthesis. In particular, if each of the a_i are the same, we define this product as a^n. That is

DEFINITION 4.21. Let $(M, \star)$ be an associative algebraic system and $a \in M$. Then

$$a^n \stackrel{\text{def}}{=} \underbrace{a \star a \star \cdots \star a}_{n \text{ times}}.$$

SOLUTION

EXERCISE 4.10. In an associative algebraic system, $(M, \star)$ prove that for any element $a \in M$ and positive integers n and m we have that

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

DEFINITION 4.22. An element $e \in M$ is *an identity* for an algebraic system $(M, \star)$ if for all $m \in M$,

$$e \star m = m \star e = m. \tag{1}$$

If such an element exists in M, we say that $(M, \star)$ *has an identity*.

Unlike associativity, it's easy to see whether or not an algebraic system has an identity by taking a glance at the multiplication table. Simply look for a row and column which are repeats of the first row and column.

EXERCISE 4.11. Show that the constant algebraic system on a set M from Exercise 4.8 has an identity iff $\#M = 1$.

SOLUTION

PROPOSITION 4.23 (Uniqueness of Identities). *If e and e' are identities for an algebraic system then $e = e'$.*

PROOF. Let $(M, \star)$ be an algebraic system with two elements e and e' which satisfy (1). Then

$$\begin{aligned} e' &= e' \star e \quad && (e' \in M \text{ and } e \text{ satisfies (1)}) \\ &= e && (e \in M \text{ and } e' \text{ satisfies (1)}). \end{aligned}$$

Therefore $e' = e$ as desired. ∎

As such, it makes sense to talk about the identity of an algebraic system, and not just an identity of an algebraic system.

WARNING! 4.24. While there is only one identity per algebraic system, several things can safely be called the identity of a single set. For instance, both $(\mathbb{Z}, +)$ and $(\mathbb{Z}, \cdot)$ are algebraic systems with identities 0 and 1 respectively, but of course $0 \neq 1$.

NOTATION 4.25. Given an algebraic system with identity, $(M, \star)$, we denote this unique identity as $e_{(M,\star)}$ or e_M (if the operation is clear).

EXAMPLE 4.26. Of course it's trivial to come up with algebraic systems which do not have an identity. Associative ones with just two elements include

$\star$	a	b
a	a	a
b	a	a

$\star$	a	b
a	a	a
b	b	b

EXAMPLE 4.27.

(1) For any set X, the algebraic systems $(\mathcal{F}_X, \circ)$, $(\mathcal{S}_X, \circ)$ each have an identity, namely the identity function $\mathbf{1}_X$.

(2) The algebraic systems $(\mathbb{H}, *)$, $(\mathbb{C}, \cdot)$, $(\mathbb{R}, \cdot)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{Z}, \cdot)$, and $(\mathbb{Z}_n, \cdot)$ each have an identity, namely 1 (in $\mathbb{Z}_n$ by 1 I mean $[\,1\,]_n$).

(3) Each $(\mathbb{H}, +)$, $(\mathbb{C}, +)$, $(\mathbb{R}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{Z}, +)$, and $(\mathbb{Z}_n, +)$ also have an identity, namely 0 (in $\mathbb{Z}_n$ by 0 I mean $[\,0\,]_n$).

(4) The identity matrix, $\mathbf{I}_n$, is an identity for $(\mathbf{M}_n(R), \cdot)$ (matrix multiplication).

(5) The zero matrix, $\mathbf{O}_n$, is an identity for $(\mathbf{M}_n(R), +)$ (matrix addition).

(6) The vector cross product does not have an identity since if $\mathbf{e}$ were the identity, then $\mathbf{e} = \mathbf{e} \times \mathbf{e} = \mathbf{0}$, so $\mathbf{e} = \mathbf{0}$, but the zero vector also can't be the identity since $\mathbf{0} \times \mathbf{i} = \mathbf{0} \neq \mathbf{i}$.

(7) The algebraic system of even numbers under multiplication, i.e. $(2\mathbb{Z}, \cdot)$, does not have an identity since it doesn't contain the number 1 (there is nothing else that can play the role of 1 since if $2 \cdot e = 2$, then subtracting 2 from both sides shows us that $2 \cdot (e - 1) = 0$. Multiplying through by $1/2$ shows us that $(e - 1) = 0$ or $e = 1$. But the number 1 is not even, it's odd.)

Definition 4.28. An algebraic system $(M, \star)$ is called a *monoid* if it

(1) has an identity,

(2) and is associative.

Exercise 4.12. Let R be a ring (e.g. $R = \mathbb{Z}_n$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$).

Solution

(1) Define

$$(a, b) \star (c, d) := (ad + bc, bd).$$

Show that $(R \times R, \star)$ is a moniod.

(2) Define

$$a * b = a + b + ab.$$

Show that $(R, *)$ is a moniod.

This exercise shows that any algebraic system can be extended into an algebraic system with identity.

Exercise 4.13. Let $(M, \star)$ be an algebraic system. Let ϵ be some object which is not in M.[II] And denote $M \cup \{\epsilon\}$ as M'. Define an operation

$$a \star' b \stackrel{\text{def}}{=} \begin{cases} a \star b & \text{whenever } a, b \in M \\ a & \text{whenever } b = \epsilon \\ b & \text{whenever } a = \epsilon\,. \end{cases}$$

Solution

(1) Prove that $(M', \star')$ is an algebraic system with identity ϵ.

(2) If $(M, \star)$ is a semigroup prove that $(M', \star')$ is a monoid.
Clarification! So the only thing to prove is that if $\star$ is associative, then so is $\star'$.

Exercise 4.14. Let n and m be integers so that $a \star b = n(a + b) + m(a \cdot b)$ is associative. What are the possibilities for n and m?

Solution

[II] Recall that this is always possible: by the axioms of set theory no set contains every object! Actually, one of the axioms say that for any set M, $\{M\} \notin M$.

Definition 4.29. Let $(M, \star)$ be an algebraic system with an identity e_M.

(1) An element $m \in M$ is an *invertible* element or a *unit* of $(M, \star)$ if $\exists\, \widehat{m} \in M$ with

$$m \star \widehat{m} = \widehat{m} \star m = e_M.$$

In which case we say that $\widehat{m}$ is *an inverse* or m.

(2) The set of all invertible elements of $(M, \star)$ is denoted $\mathcal{U}(M, \star)$.[III]

(3) If $\mathcal{U}(M, \star) = M$, we say that $(M, \star)$ is a *loop*.

That is an algebraic system is a loop iff every element is invertible. Thus, given an algebraic system $(M, \star)$ with identity e_M

$$\mathcal{U}(M, \star) = \{m \in M : \exists \widehat{m} \in M \text{ s.t. } m \star \widehat{m} = \widehat{m} \star m = e\}.$$

Another way of thinking of this is as the requirement that the equations

$$x \star m = e_{(M,\star)} \quad \text{and} \quad m \star x = e_{(M,\star)}$$

have a (common) solution.

Remark 4.30.

(1) The relation that b is an inverse to a is symmetric in the sense that a is an inverse to b iff b is an inverse to a.

(2) Of course the notation of an inverse does not make any sense if the algebraic system does not have an identity in the first place!

(3) One could say that b is a *right inverse* of a and, dually, that a is a *left inverse* of b if $a \star b = e_M$. Since we have so much freedom, it's trivial to come up with examples like this. For instance, if $M = \{e, a, b\}$ with binary operation

$\star$	e	a	b
e	e	a	b
a	a	b	e
b	b	a	e

then a is a left inverse to b, but b is not a left inverse of a. Meanwhile b is a (2-sided) inverse of itself. We can even find examples of associative algebraic systems which allow for such pathologies, see Remark 4.35 below. We have little use of these definitions.

Example 4.31. In the algebraic system given in Example 4.11 the identity of M is e, and every element has an inverse. (Thus the algebraic system is a loop.) Namely an inverse of e is e itself, the element a actually has two inverses, namely a itself and b, and the element b has a as an inverse.

Example 4.32. All of the examples of algebraic systems with an identity listed in the above example, Example 4.27, are monoids. The algebraic system $(\mathbb{N}, \star)$ with $a \star b = |a - b|$ (this distance between a and b on the real line) has an identity, namely 0, and every element has an inverse; for any element a the inverse of a is a itself since

$$a \star a = |a - a| = |0| = 0 = e_{(N,\star)}.$$

But it is not a monoid since it is not associative. Indeed,

$$(3 \star 2) \star 1 = 1 \star 1 = 0 \neq 3 \star (2 \star 1) = 3 \star 1 = 2.$$

The algebraic system in 4.11 is a loop (actually both a and b are inverses to a) but is not associative.[IV] This is because

$$(a \star b) \star b = e \star b = b \neq e = a \star b = a \star (b \star b).$$

Neither algebraic system from example 4.26 is a monoid since neither has an identity. (However both are associative.)

We spent a lot of time in the last chapter talking about which elements have inverses in various contexts. In definition 3.135 we defined what it means for an element to be invertible in a ring. Since the multiplicative part (as well as the additive part) of any ring R is a monoid, our two meanings of $\mathcal{U}(R)$ correspond. Note that we never really need to talk about $\mathcal{U}(R, +)$ since every element has a negative and hence this set is simply equal to R itself. In particular, for the monoid $(\mathbb{Z}_n, \cdot)$ the definition given of $\mathcal{U}(\mathbb{Z}_n)$ (which always means $\mathcal{U}(\mathbb{Z}_n, \cdot)$ in our context) given in this chapter and the last are identical.

Example 4.33. Again, whether or not an element of an algebraic system is invertible depends on the set and the operation. For instance, in $(\mathbb{Z}, +)$ every element is invertible. In fact the additive part of every ring, $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$, $(\mathbb{H}, +)$, and $(\mathbb{Z}_n, +)$. The inverse being negative the element.

On the other hand, in the algebraic system $(\mathbb{N}, +)$ only the number 0 is invertible. This is because for every non-zero number has a negative negative (don't you just love using the same word twice in a row?) Every non-zero element of $(\mathbb{Q}, \cdot)$ is invertible and the only non-zero elements of $(\mathbb{Z}, \cdot)$ which are invertible are ± 1. The invertible elements of $\mathbf{M}_n(R)$ are, by Theorem 3.231, $\mathbf{GL}_n(R)$. The invertible functions on a set X, $\mathcal{F}_X$, are precisely the bijections, $\mathcal{S}_X$. This is due to Theorem 2.53.

We make a couple of cute tables. We have the trivial, boring examples which all have addition as their operation:

$(M, \star)$	$(\mathbb{Z}, +)$	$(\mathbb{Q}, +)$	$(\mathbb{R}, +)$	$(\mathbb{C}, +)$	$(\mathbb{H}, +)$	$(\mathbb{Z}_n, +)$
$\mathcal{U}(M)$	$\mathbb{Z}$	$\mathbb{Q}$	$(\mathbb{R}, +)$	$\mathbb{C}$	$\mathbb{H}$	$\mathbb{Z}_n$

Note in this string of examples we could have included direct products of these, i.e. $(\mathbb{R}^n, +)$ as well as matrices (under addition), i.e. $(\mathbf{M}_n(R), +)$ where R is any commutative ring. The more interesting examples come when we consider an operation which is not addition[V]

$(M, \star)$	$(\mathbb{Z}, \cdot)$	$(\mathbb{Q}, \cdot)$	$(\mathbb{R}, \cdot)$	$(\mathbb{C}, \cdot)$	$(\mathbb{H}, *)$	$(\mathbf{M}_n(R), \cdot)$	$(\mathcal{F}_X, \circ)$	$(\mathbb{Z}_n, \cdot)$
$\mathcal{U}(M)$	$\{\pm 1\}$	$\mathbb{Q}^\times$	$\mathbb{R}^\times$	$\mathbb{C}^\times$	$\mathbb{H}^\times$	$\mathbf{GL}_n(R)$	$\mathcal{S}_X$	$\mathcal{U}(\mathbb{Z}_n)$

(Of course the last entry says that $\mathcal{U}(\mathbb{Z}_n) = \mathcal{U}(\mathbb{Z}_n)$ which seems rather bland. But of course we have seen in Theorem 3.139 that

$$\mathcal{U}(\mathbb{Z}_n) = \{[a]_n : \gcd(a, n) = 1\}.)$$

The additive parts of rings

[IV] We could see this instantly by using 4.34 below.

[V] Really I should say, "which is not called addition".

THEOREM 4.34 (Uniqueness of Inverses). *Suppose that $(M, \star)$ is a monoid. Suppose that $h_1, h_2, g \in M$ and that h_1 and h_2 are inverses for g. Then $h_1 = h_2$.*

PROOF. (We will actually show that if h_1 is a left inverse of g and h_2 is a right inverse of g [both defined in the obvious way] that $h_1 = h_2$.) We have

$$\begin{aligned} h_1 &= h_1 \star e_M && \text{(Def'n of } e_M\text{.)} \\ &= h_1 \star (g \star h_2) && \text{(Since } h_2 \text{ is an inverse of } g\text{.)} \\ &= (h_1 \star g) \star h_2 && \text{(Since monoids are associative.)} \\ &= e_M \star h_2 && \text{(Since } h_1 \text{ is an inverse of } g\text{.)} \\ &= h_2. && \text{(Def'n of } e_M\text{.)} \end{aligned}$$

In conclusion $h_1 = h_2$ as desired. ■

REMARK 4.35. In **??** above we gave a non-associative example of an algebraic system with only one-sided inverses. In this example we give an associative algebraic system with an infinite number of right inverses, but without any left inverses. For any real number k we define shift maps

$$\mathcal{S}_k : (a_1, a_2, a_3, a_4, \ldots) \mapsto (k, a_1, a_2, a_3, \ldots) \qquad \text{(Shift Right)}$$

and

$$\mathbf{K} : (b_1, b_2, b_3, b_4, \ldots) \mapsto (b_2, b_3, b_4, \ldots). \qquad \text{(Shift Left)}$$

Note that we have for all $k \in \mathbb{R}$, $\mathbf{K} \circ \mathcal{S}_k$ is the identity.[VI] Hence we have an infinite number of right inverses of $\mathbf{K}$ (one for every real number) in the monoid of functions on sequences $(\mathcal{F}_{\mathbb{R}^\infty}, \circ)$. Also, note that no $\mathcal{S}_k$ is a left inverse of $\mathbf{K}$. To see this select a real number, r, not equal to k. Then

$$\begin{aligned} (\mathcal{S}_k \circ \mathbf{K})(r, 0, 0, 0, \ldots) &= \mathbf{K}(r, 0, 0, 0, \ldots) \\ &= \mathcal{S}_k(0, 0, 0, 0, \ldots) \\ &= (k, 0, 0, 0, \ldots).^{\text{VII}} \end{aligned}$$

And since $k \neq r$ we have that $(r, 0, 0, 0, \ldots) \neq (k, 0, 0, 0, \ldots)$ and so $\mathcal{S}_k \circ \mathbf{K} \neq \mathbf{1}_{\mathbb{R}^\infty}$.

Luckily, in most familiar situations such pathological behavior (i.e. nonsense) does not occur. For instance, for matrices, or for maps between finite sets (e.g. elements of $\mathcal{F}_n$) we have a right inverse for something iff we have a left inverse for it and they coincide.

Therefore in a monoid we can make the following definition

DEFINITION 4.36. If $(M, \star)$ is monoid and $g \in M$ is invertible, then we denote the one and only inverse of g as g^{-1}.[VIII]

THEOREM 4.37 (Socks and Shoes). *If a and b are invertible elements of a monoid $(M, \star)$ (i.e. are elements of $\mathcal{U}(M, \star)$), then $a \star b$ is also invertible (hence also an element of $\mathcal{U}(M, \star)$) with inverse $b^{-1} \star a^{-1}$. (Hence $\mathcal{U}(M, \star)$ is closed under $\star$ and is therefore a sub-algebraic system of $(M, \star)$.)*

Proof. Since we are in a monoid, we have unique of inverses, so we just need to verify that $b^{-1} \star a^{-1}$ is *an* inverse of $a \star b$ to conclude that it is the inverse of $a * b$. Indeed, this is easy since

$$\begin{aligned}
(a \star b) \star (b^{-1} \star a^{-1}) &= a \star b \star b^{-1} \star a^{-1} && \text{(By Theorem 4.20)} \\
&= a \star e_M \star a^{-1} && \text{(Def'n of } b^{-1}\text{)} \\
&= (a \star e_M) \star a^{-1} && \text{(Putting in paren for emph)} \\
&= a \star a^{-1} && \text{(Def'n of } e_M\text{)} \\
&= e_M. && \text{(Def'n of } a^{-1}\text{).}
\end{aligned}$$

The proof that $(b^{-1} \star a^{-1}) \star (a \star b)$ is also the identity is similar. ∎

The reason for the name of this theorem is quite a good one. If you put your socks and then your shoes on in the morning how would you reverse this process? Clearly you wouldn't remove your socks and then your shoes. You would reverse things in the *opposite* order ... you would first take off your shoes and then your socks!

Here is the most important definition in the course.

Loop + Monoid = Group

Definition 4.38. An algebraic system is a *group* if it is both a monoid and a loop.

Starting in the next chapter, we will nearly solely focus on groups. At the beginning of the next chapter we will expand the definition of a group out a bit. For now we will contend ourselves with the following, which is a corollary of the above.

Corollary 4.39. *For any monoid* $(M, \star)$, $\mathcal{U}(M, \star)$ *is a group.*

Proof. Let's write down the definition

$$\mathcal{U}(M, \star) = \{m \in M : \exists \widehat{m} \in M \text{ s.t. } m \star \widehat{m} = \widehat{m} \star m = e\}.$$

again for quick reference.

(1) By Socks and Shoes, Theorem 4.37, $\mathcal{U}(M, \star)$ is a sub-algebraic system. In more detail we know that if $a, b \in \mathcal{U}(M)$, then $a \star b \in \mathcal{U}(M)$ since there exists an element g (namely $g = b^{-1} \star a^{-1}$) so that $(a \star b) \star g = g \star (a \star b) = e_M$.

(2) By Theorem 4.18 it is a semi-group. (It is a subset of an associative algebraic system which is closed under the operation.)

(3) Since $(M, \star)$ is assumed to be a monoid, there exists an identity, $e_M \in M$. This identity is invertible since

$$e_M \star e_M = e_M = e_M \star e_M \text{[IX]}$$

Therefore, by definition $e_M \in \mathcal{U}(M)$ Clearly e_M is an identity of the algebraic system $\mathcal{U}(M)$ since for all $m \in \mathcal{U}(M)$, $m \in M$ so

$$m \star e_M = e_M \star m = m.$$

(4) Let $a \in \mathcal{U}(M)$ be arbitrary. We must show that exists a $b \in M$ so that

(a) $b \in \mathcal{U}(M)$

(b) $a \star b = b \star a = e_M$.

Can solve $x \star e_M = e_M \star x = e_M$ w $x = e_M$.

The obvious candidate is to take $b := a^{-1}$. This makes sense since by definition $a \in \mathcal{U}(M)$ means there exists an $a^{-1} \in M$. Note that this $b = a^{-1}$ satisfies the second condition. The only subtlety is this we need b to also satisfy the first condition of the two just listed.[X] What do we need for $b = a^{-1} \in \mathcal{U}(M)$? Well we need an element

[IX] Yes, it looks silly writing the same thing twice, but in the second $e_M \star e_M$ "the order has been switched."

[X] Indeed, this is quite subtle and I think just about everybody misses this the first time they are looking at it!

c so that $c \star b = b \star c = e_M$. But of course $c = a$ satisfies this, and we have proven that $b \in \mathcal{U}(M)$. Therefore b satisfies both conditions that were necessary. ■

Exercise 4.15. (Generalized Socks and Shoes) Suppose that $a_1, \ldots, a_n$ are invertible elements of a monoid. Show that

$$a_1 \star \cdots \star a_n$$

is invertible with inverse

$$a_n^{-1} \star \cdots \star a_1^{-1}.$$

SOLUTION

In this exercise we briefly examine a different product (that I hope you are asking yourself about) defined on $\mathbb{H}$ and show that it rather sucks.
Define a product $\divideontimes$ on $\mathbb{H}$ via[XI]

$$\mathsf{q} \divideontimes \mathsf{p} := (q_R \cdot p_R + \mathbf{q}_P \bullet \mathbf{p}_P, q_R \cdot \mathbf{p}_P + q_R \cdot \mathbf{p}_P + \mathbf{q}_P \times \mathbf{p}_P).$$

That is $\divideontimes$ is the same as the usual quaternion multiplication $*$ except the sign in front of the dot product is made positive. In particular, we still have for any $c \in \mathbb{R}$, $c \cdot \mathsf{q} = c \divideontimes \mathsf{q}$ and for perpendicular pure quaternions $\mathbf{v}$ and $\mathbf{w}$, $\mathbf{v} \divideontimes \mathbf{w} = \mathbf{v} \times \mathbf{w}$. The one big difference comes from the fact that we now have for any pure quaternion $\mathbf{v}$, $\mathbf{v} \divideontimes \mathbf{v} = \|\mathbf{v}\|^2$, the length squared, instead of negative the length squared. In particular,

$$\mathbf{i} \divideontimes \mathbf{i} = \mathbf{j} \divideontimes \mathbf{j} = \mathbf{k} \divideontimes \mathbf{k} = 1.$$

Another definition you'll need in this problem is that of the light cone: it's the set of $a + \mathbf{v} \in \mathbb{H}$ with

$$|a| = \|\mathbf{v}\|.$$

Exercise 4.16.

SOLUTION

(1) Prove that $(\mathbb{H}^\times, \divideontimes)$ has an identity.

(2) Prove that $(\mathbb{H}^\times, \divideontimes)$ is not associative.
Hint! Start simple. If you try something, anything, you are likely to succeed!

(3) Prove that $(\mathbb{H}^\times, \divideontimes)$ is *first coordinate associative* in the sense that for all q, p, and r, we have that

$$\mathbf{Re}((\mathsf{q} \divideontimes \mathsf{p}) \divideontimes \mathsf{r}) = \mathbf{Re}(\mathsf{q} \divideontimes (\mathsf{p} \divideontimes \mathsf{r})).$$

(4) Prove that an element of $\mathsf{q} \in \mathbb{H}$ is invertible under $\divideontimes$ iff q is not in the light cone.
Hint! If q is in the light cone, prove that q is a zero divisor. Then prove using first coordinate associativity that q cannot have an inverse. On the other hand, if q is not in the light cone prove that q is proportional to an inverse.

(5) Prove that an element of $\mathsf{q} \in \mathbb{H}$ is invertible under $\divideontimes$, then the inverse is unique.
Careful! Do not use associativity!

[XI] This definition is not used anywhere else in the book. The purpose of this exercise is **not** to get you to understand $\divideontimes$, but it **is** to get you appreciate the nice properties of $*$ a bit more.

Definition 4.40. Two elements m and m' in an algebraic system $(M, \star)$ *commute* if $m \star m' = m' \star m$. $(M, \star)$ is called *abelian* (or *commutative*) if any two elements of M commute.

In symbols, $(M, \star)$ is abelian if

$$\forall m, m' \in M,\ m \star m' = m' \star m.$$

Example 4.41.

(1) In any algebraic system any element commutes with itself.

(2) By definition of an identity, if an algebraic system contains an identity, then every element of that same algebraic system commutes with that identity.

(3) In an associative algebraic system, Exercise 4.10, tells us that any two powers of a single element commute with each other. In particular, for any element a in an algebraic system and any positive integer n, we have that a commutes with a^n.

(4) Also by definition, any element commutes with any inverse of itself. This is because for a and an inverse $\widehat{a}$ of a, both products

$$a \star \widehat{a} \quad \text{and} \quad \widehat{a} \star a$$

equal the identity, and hence are equal to each other.

Remark 4.42. Students have a hard time getting used to the fact that things do not commute. We have already seen several situations where things do not commute. Here's a giant list of things designed to completely and utterly destroy any shred of hope in your mind that "things should commute."

(1) In high school one learns "order of operations." (By now this is something that must students do not really even thing about anymore.) For instance, in $4 \cdot x + 5$ we all know by now that we first multiply by 4, then we add 5. The complexity of this comes from the fact that operations of multiplication and addition do not commute: indeed if we added 5 first and then multiplied by 4 we'd end up with $4 \cdot (x + 5) = 4 \cdot x + 20$ which never coincides with the original order since $5 \neq 15$. Similarly taking powers does not commute with multiplication (or addition), e.g. $3 \cdot 5^2 \neq (3 \cdot 5)^2$. Since these operations do not commute one is forced to impose this (arbitrary) order of operations (it essentially amounts doing the most complex operations first, but this explanation would apparently save too much time in high school, and save an excuse to memorize a cool acronym ⟨what was that acronym? PDNMSA ... **P** LEASE **D**O **N**OT **M**EMORIZE **S**TUPID **A**CRONYMS ⟩ ... !)

(2) The quaternions $\mathbf{i}$ and $\mathbf{j}$ do not commute ($\mathbf{ij} = \mathbf{k}$ but $\mathbf{ji} = -\mathbf{k}$). Indeed Theorem 3.75 says that two quaternions commute iff their pure parts are parallel.

(3) Most pairs of matrices also do not commute. For instance, the matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

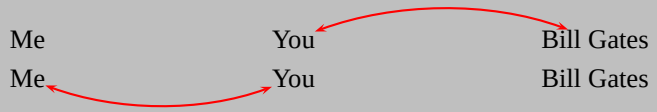
never commute!

(4) A more formal way of stating **(1)** is in terms on non-commutativity of certain functions. For instance the functions $f(x) = x^2$ and $g(x) = 3 \cdot x$ do not commute we nearly summed up by the statement that $3 \cdot 5^2 \neq (3 \cdot 5)^2$ we made there.

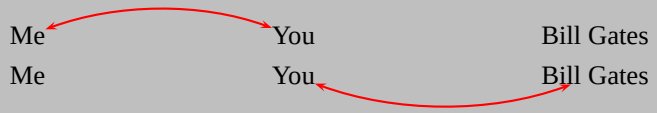
Indeed, most pairs of functions do not commute. This is often the bane of many Calculus I students. For instance, I would be a rich man if I had a dollar for every Calculus I student who made the mistake of thinking that $\sin(x^2)$ was equal to $(\sin x)^2$ (i.e. they are assuming that the functions $\sin x$ and x^2[XII] commute.)

(5) I have also already pointed to an everyday experience of socks and shoes which do not commute. It is easy to come up with many examples of this sort of phenomena: nearly whenever to actions "depend" on each other. For instance pouring soda and taking off the cap. On the other hand, when actions do not "depend" on each other they tend to commute: it probably does not matter if you wash your hair first or brush your teeth.

(6) Permutations also rarely commute. For instance, if me, you, and Bill Gates were all standing in a straight line with their daily pay checks and if you and Bill swapped and then me and you swapped



I would end up with Bill's cash (he would end up with yours and you would end up with mine), but if me and you swapped first and then you and Bill swapped *you* would end up with



Bill's cash (he would get mine, and I would get yours). These two permutations lead to very different situations! More formally we just gave a aesthetic reason why the permutations

$$\begin{pmatrix} \text{Me} & \text{You} & \text{Bill Gates} \\ \text{You} & \text{Me} & \text{Bill Gates} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \text{Me} & \text{You} & \text{Bill Gates} \\ \text{Me} & \text{Bill Gates} & \text{You} \end{pmatrix}$$

do not commute.

(7) (**Physics Example**) In physics, the lack of comutivity gives rise to the famous *Heisenberg Uncertainty Principle*. Indeed, in a quantum mechanics class, up to a constant, the uncertainty is proportional to the difference

$$\widehat{x} \circ \widehat{p} - \widehat{p} \circ \widehat{x},$$

where $\widehat{x}$ is the quantum mechanical operator which measures position along the x direction (it's just multiplication by x) and $\widehat{p}$ is the quantum mechanical operator which measures momentum (in one dimension, it is the operator $\frac{h}{2\pi i}\frac{\partial}{\partial x}$). Of course this difference would be 0 if $\widehat{x}$ and $\widehat{p}$ commuted. Intuitively, if we first measure the position of a particle, then this measurement changes what we'd get when we measure the momentum of the particle. As I understand it, the reason why this happens is that the very act of measuring a particles position requires nailing the particle with something (e.g. a photon), which will cause the particle to change the way it was moving.

EXERCISE 4.17. Show that every monoid with two elements is abelian.

SOLUTION

[XII] Really this functions shouldn't be called $\sin x$ and x^2, they should be thought of as the functions which respectively takes the sin of a number and takes a number and squares it.

EXERCISE 4.18. In this exercise we find all associative algebraic systems on two elements.

SOLUTION

(1) Show that the only two associative, non-abelian algebraic systems on the set $\{a, b\}$ are given by the following tables:

$*_1$	a	b
a	a	a
b	b	b

and

$*_2$	a	b
a	a	b
b	a	b

Hint! Show that if $a^2 = b$, then a and b must commute by considering the products $a \star (a \star a)$ and $(a \star a) \star a$.

(2) Show that the only monoids on the set $\{a, b\}$ are given by one of the following four tables:

$\star_1$	a	b
a	a	b
b	b	a

$\star_2$	a	b
a	a	a
b	a	b

$\star_3$	a	b
a	b	a
b	a	b

$\star_4$	a	b
a	a	b
b	b	b

Hint! Perhaps Exercise 4.17 would be helpful.

(3) Show that the only two associative, abelian algebraic systems without an identity are given by

$\bullet_1$	a	b
a	a	a
b	a	a

and

$\bullet_2$	a	b
a	b	b
b	b	b

(4) Use the previous parts to conclude that there are exactly 8 associative algebraic systems which can be defined on a set with two elements.

EXERCISE 4.19.

SOLUTION

(1) Consider the binary operation $\bullet$ defined on $\mathbb{N}$ via $a \bullet b = \gcd(a, b)$. Show that $(\mathbb{N}, \bullet)$ is associative but does not have an identity.

(2) Instead, consider the binary operation $\circ$ defined on $\mathbb{N}$ via $a \circ b = \operatorname{lcm}(a, b)$. Show that $(\mathbb{N}, \circ)$ is a monoid, but only the identity is invertible.

Definition 4.43. Let $*$ and $\star$ be two binary operation defined on the set same set M. We say that $\bullet$ distributes over $*$ if

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

(3) Continuing the problem

(a) Prove that gcd distributes over lcm;

(b) Prove that lcm distributes over gcd;

(c) Where have you seen this "double distributivity" before?

Exercise 4.20. Two algebraic systems $(\{a, b\}, \star)$ and $(\{a, b\}, \bullet)$ are *equivalent* (denoted $\star \sim \bullet$) if $\star = \bullet$ or if one can swap every occurrence the letters a and b in the first table (including on the top) to get the second (by this I mean the same multiplication, clearly the top of the table will be written in a different order). For instance, the algebraic system determined by the table

$\star$	a	b
a	b	a
b	a	b

is equivalent to itself and to the algebraic system determined by

$\star$	b	a
b	a	b
a	b	a

which we write (only to satisfy our NEED to write a before b!) as

$\star$	a	b
a	a	b
b	b	a

(1) Show that $\sim$ is an equivalence relation.

(2) Suppose that $\star \sim \bullet$. Prove that

- If $(\{a, b\}, \star)$ is associative, then so is $(\{a, b\}, \bullet)$.
- If $(\{a, b\}, \star)$ has an identity, then so does $(\{a, b\}, \bullet)$.
- If $(\{a, b\}, \star)$ is a monoid, then so is $(\{a, b\}, \bullet)$.

Solution

(3) Which of the tables from Exercise 4.18 are equivalent?

(4) Up to equivalence, how many associative algebraic systems are there on a set with two elements?

EXERCISE 4.21. Construct a non-abelian monoid on the set $\{e, a, b\}$.

SOLUTION

II Basics of Group Theory

5 Group Basics

This is a course of group theory, but we have yet to define what a group is. That is the point of this chapter. Also in this chapter we define a few of the important things we will need to understand and distinguish between different groups including the order of a group and the order of an element of a group.

5.1 First Definitions and Examples

Recall that if $(M, \star)$ is a monoid, we defined the set of invertible elements of $(M, \star)$

$$\mathcal{U}(M, \star) \text{ or more sloppily } \mathcal{U}(M) \stackrel{\text{def}}{=} \{ m \in M : \exists n \in M \text{ s.t. } m \star n = e_M = n \star m \}.$$

We repeat the following definition, which was discussed briefly in the last chapter, of the most important definition in the course.

Definition 5.1. An algebraic system $(G, \star)$ is a *group* if

(1) $(G, \star)$ is a monoid;

(2) $\mathcal{U}(G, \star) = G$.

We say that G forms a group under $\star$.

The first condition says that $\star$ is an associative product with an identity, and the second condition says that every element of G has an inverse. That is

$$\forall g \in G, \exists \widehat{g} \in G \text{ s.t. } g \star \widehat{g} = \widehat{g} \star g = e_{G,\star}.$$

Since we have shown in Theorem 4.34 that this $\widehat{g}$ is unique, we safely denote for any element g of a group, the unique inverse of g as g^{-1}.

Warning! 5.2. We again repeat the theme of Example 4.33. We require that the inverse of an element must live in the same set! For instance, the algebraic system $(\mathbb{N}, +)$ is a monoid, but it is not a group since 1 does not have inverse (there does not exist an element of $a \in \mathbb{N}$ so that $1 + a = 0 = e_{\mathbb{N},+}$). There does exist an element in $\mathbb{Z}$ which solves this, but by definition, the inverse must live in $\mathbb{N}$ for it to count! Similarly $(\mathbb{Z}^\times, \cdot)$ is not a group since 2 is not invertible in $\mathbb{Z}$ (and again, we don't give a crap that $1/2$ lives in the larger set $\mathbb{Q}^\times$.)

DEFINITION 5.3. A group is *abelian* if the algebraic system is abelian.[I]

So in symbols, a group $(G, \star)$ is abelian iff for all $g_1, g_2 \in G$ we have

$$g_1 \star g_2 = g_2 \star g_1.$$

EXAMPLE 5.4. For any monoid $(M, \star)$, Corollary 4.39 $(\mathcal{U}(M), \star)$ is a group. In particular all sets in the bottom row of the table in Example 4.33 are groups with the induced operations. Repeating these tables here, we have both the boring, additive examples where everything is the same listed on the top row, and the more interesting examples listed on the bottom row:

Groups →	$(M, \star)$	$(\mathbb{Z}, +)$	$(\mathbb{Q}, +)$	$(\mathbb{R}, +)$	$(\mathbb{C}, +)$	$(\mathbb{H}, +)$	$(\mathbf{M}_n(R), +)$	$(\mathbb{Z}_n, +)$	
Groups →	$(\mathcal{U}(M), \star)$	$(\mathbb{Z}, +)$	$(\mathbb{Q}, +)$	$(\mathbb{R}, +)$	$(\mathbb{C}, +)$	$(\mathbb{H}, +)$	$(\mathbf{M}_n(R), +)$	$(\mathbb{Z}_n, +)$	
Not Groups →	$(M, \star)$	$(\mathbb{Z}, \cdot)$	$(\mathbb{Q}, \cdot)$	$(\mathbb{R}, \cdot)$	$(\mathbb{C}, \cdot)$	$(\mathbb{H}, *)$	$(\mathbf{M}_n(R), \cdot)$	$(\mathbb{Z}_n, \cdot)$	$(\mathcal{F}_n, \circ)$
Groups →	$(\mathcal{U}(M), \star)$	$(\{\pm 1\}, \cdot)$	$(\mathbb{Q}^\times, \cdot)$	$(\mathbb{R}^\times, \cdot)$	$(\mathbb{C}^\times, \cdot)$	$(\mathbb{H}^\times, *)$	$(\mathbf{GL}_n(R), \cdot)$	$(\mathcal{U}(\mathbb{Z}_n), \cdot)$	$(\mathcal{S}_n, \circ)$

(Again, in general, there is no additive operation one can define on $\mathcal{S}_X$.) This gives us a running start of 17 examples. Nearly every example of a group we encounter in this text will contained inside one of these 17 examples (a few, however, will be equivalence classes in a subset of one of these 17 examples).

We begin to carve out some interesting examples from the gang of 17 listed above. To unload definition, let's expand it as much as possible into four axioms:

A set G with a binary operation $\star$ is a group if $\forall g, g', g'' \in G$

Closure $g \star g' \in G$.

Assoc. $(g \star g') \star g'' = g \star (g' \star g'')$.

Identity $\exists e_G \in G$ so that $e_G \star g = g = g \star e_G$.

Inverses $\exists \widehat{g} \in G$ so that $g \star \widehat{g} = e_G = \widehat{g} \star g$.

The first example we give of an example that isn't one of the gang of 17 listed in previous example is $\mathbb{S}^3$. Recall that it is, by definition, the subset of quaternions of length 1, i.e.

$$\mathbb{S}^3 = \{\mathsf{q} \in \mathbb{H} : \|\mathsf{q}\| = 1\}.$$

[I] For historic reasons we always say that a group is abelian, and we never say that a group is *commutative*. Meanwhile the opposite is true for a ring. That is we never call a ring *abelian* we always say that a ring is commutative.

Example /Theorem 5.5. $(\mathbb{S}^3, *)$ is a non-abelian group.

Proof. We verify that $\mathbb{S}^3$ satisfies the four parts of the definition:

Closure The (quaternion) product of any two elements of $\mathbb{S}^3$ is an element of $\mathbb{S}^3$ by Theorem 3.101**(1)**.

Assoc. The associativity of the quaternion product on $\mathbb{S}^3$ is inherited from the associativity of the quaternion product which holds for all of $\mathbb{H}$. (See Theorem 4.18 and Theorem 3.98).

Identity The quaternion 1 has $\| 1 \| = 1$, and is therefore an element of $\mathbb{S}^3$. Since 1 satisfies

$$1 * q = q * 1 = q$$

for all $q \in \mathbb{S}^3$, 1 is an identity for $(\mathbb{S}^3, *)$.

Inverses The quaternion inverse of any element of $\mathbb{S}^3$ is an element of $\mathbb{S}^3$ by Theorem 3.101**(2)**. Therefore, for any element $q \in \mathbb{S}^3$, there exists an element $q^{-1} \in \mathbb{S}^3$ so that

$$q * q^{-1} = q^{-1} * q = 1 = e_{(\mathbb{S}^3, *)}.$$

Moreover, this group is not abelian since **i** and **j** are elements of $\mathbb{S}^3$ which do not commute. ■

Also note the exact same proof would prove that $\mathbb{S}^1$ which is the set

$$\mathbb{S}^1 = \{z \in \mathbb{C} : \| z \| = 1\}$$

is also a group. However $\mathbb{S}^1$, unlike $\mathbb{S}^3$, **is** abelian. This is because every element of $\mathbb{S}^1$ is a complex number and the product of complex numbers, unlike the product of quaternions, is commutative.

Of course $(\mathbb{S}^3, +)$ is not a group since $\mathbb{S}^3$ is not even closed under $+$. For instance, $\mathbf{i} + \mathbf{i} = 2\mathbf{i}$ has length 2 and, as such, is not an element of $\mathbb{S}^3$ (which are by definition, quaternions of length 1).

Exercise 5.1. For any positive integer, prove that $\{\{n\}\}$ is a group. (See Definition 3.56 back on page 97.)

SOLUTION

Exercise 5.2. Let R be a commutative ring with more than one element[II] and let $n \in \mathbb{Z}^+$. Prove

(1) $\mathbf{SL}_n(R)$ is a group;

(2) $\mathbf{SL}_n(R)$ is non-abelian.
Careful! Make sure you choose general enough matrices that make sense for any ring R.

SOLUTION

Exercise 5.3. Let R be one of the rings, $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$. Consider the set of formal[III] functions of the form

$$f_{a,b,c,d}(z) = \frac{az+b}{cz+d}$$

where $a, b, c, d \in R$ which we denote as $\mathbf{PM}_2(R)$.[IV]

Solution

(1) Show that $f_{a,b,c,d}$ is constant iff $ad - bc = 0$.

(2) Show that $\mathbf{PM}_2(R)$ is closed under composition, and hence $(\mathbf{PM}_2(R), \circ)$ is an algebraic system.

(3) Show that the function $f_{1,0,0,1}$ (which is the function $f_{1,0,0,1}(z) = z$) is an identity for $(\mathbf{PM}_2(R), \circ)$.

(4) Finish the proof that $(\mathbf{PM}_2(R), \circ)$ is a monoid by proving associativity.
Hint! This is trivial!

(5) Suppose that $f_{a,b,c,d}$ is not constant. Prove that $f^{-1}_{a,b,c,d}$ is $f_{d,-b,-c,a}$ by proving that

$$f_{a,b,c,d} \circ f_{d,-b,-c,a}(z) = z = f_{d,-b,-c,a} \circ f_{a,b,c,d}(z).$$

Conclude that $\mathcal{U}(\mathbf{PM}_2(R), \circ)$ is the set of non-constant functions in $(\mathbf{PM}_2(R), \circ)$
Hint! Clearly a constant function is not invertible, so $\mathcal{U}(\mathbf{PM}_2(R), \circ)$ is contained inside of the set of non-constant functions in $\mathbf{PM}_2(R)$ and this part shows that the reverse inclusion is also true.

(6) Conclude that

$$\mathbf{PGL}_2(R) \stackrel{\text{def}}{=} \left\{ \frac{az+b}{cz+d} : ad - bc \neq 0 \right\}$$

forms a group under composition. (Called the *Projective Linear Group* with coefficients in R.)

Exercise 5.4. Let X be any set and suppose that A and B are subsets of X. The *symmetric difference* of A and B is the set

Solution

$$A \,\Delta\, B \stackrel{\text{def}}{=} (A \setminus B) \cup (B \setminus A).$$

Show that

$$(\mathcal{P}(X), \Delta)$$

forms an abelian group.
Hint! It is surprisingly difficult to prove that this operation is associative.

Exercise 5.5. Let $\star$ be the binary operation from Exercise 4.12 on page 224. Show that

Solution

$$(\mathbb{R} \setminus \{-1\}, \star)$$

is an abelian group.

EXERCISE 5.6. Write down a multiplication table for the 8 matrices

$$\left\{\begin{pmatrix} [1]_2 & [a]_2 & [b]_2 \\ [0]_2 & [1]_2 & [c]_2 \\ [0]_2 & [0]_2 & [1]_2 \end{pmatrix} : [a]_2, [b]_2, [c]_2 \in \mathbb{Z}_2\right\}.$$

SOLUTION

Argue directly from the table that this determines a non-abelian group.[v] This group will be denoted $UT_3(\mathbb{Z}_2)$.
Hint! Since these are matrices you do not need to prove that the product is associative.

DEFINITION 5.6. Given a group $(G, \star)$, the *order* of $(G, \star)$, denoted as $|G|$ or $|(G, \star)|$, is the number of elements in G. That is

$$|(G, \star)| = \#G.$$

If there are an infinite number of elements in the set G we say that the group $(G, \star)$ has *infinite order* and we write

$$|(G, \star)| = \infty.$$

For example, we have that $|(\mathbb{Z}_n, +)| = n$ (by Corollary 3.125) and $|(\mathcal{S}_n, \circ)| = n!$ (by Theorem 2.60). For any prime number, we will see that $\left(\mathbf{SL}_2\left(\mathbb{Z}_p\right), \cdot\right)$ has order $(p+1)\,p\,(p-1)$ (we will also come up with more general formulas for p prime and $n > 2$). The group $\mathcal{U}(\mathbb{Z}_n)$ is a group of order $\vartheta(n)$ where ϑ is the Euler totient function defined in Remark 3.4.3 on page 153. Also we have that $(\mathbb{S}^3, *)$, $(\mathbb{C}^\times, \cdot)$, $(\mathbb{H}^\times, *)$, $(\mathbb{S}^1, \cdot)$, and $(\mathbb{Z}, +)$ all have infinite order.

Let's make a table: (We update this table throughout the text. The finished product appears on page 863.)

Set	$\mathbb{Z}_n$	$\mathcal{S}_n$	$\mathbf{SL}_2\left(\mathbb{Z}_p\right)$	$\mathcal{U}(\mathbb{Z}_n)$	$\mathbb{H}^\times$	$\mathbb{S}^3$	$\mathbb{C}^\times$	$\mathbb{S}^1$	$\{\{n\}\}$	$\mathbb{Z}$
Operation	$+$	$\circ$	$\cdot$	$\cdot$	$*$	$*$	$\cdot$	$\cdot$	$\cdot$	$+$
Order	n	$n!$	$(p^3 - p)$	$\vartheta(n)$	∞	∞	∞	∞	n	∞
Abelian?	yes	no	no	yes	no	no	yes	yes	yes	yes

5.2 SYMMETRY GROUPS

We have covered all of the families of groups we will be using as examples in the previous section with the exception of a couple which are all related to the set of symmetries of an object. Every object in the universe has a certain set of symmetries to it. An interesting class of groups come to us as the set of symmetries of an object. In fact, group theory is often called the study of symmetry.

Suppose I had a soccer ball (see Figure 5.1) on my desk, which had no distinctive labels on it, and some student moved the soccer ball in a certain way while I was not looking. Moreover, suppose the sneaky student moved it in such a way that when I looked back I could not tell the soccer ball had been touched.

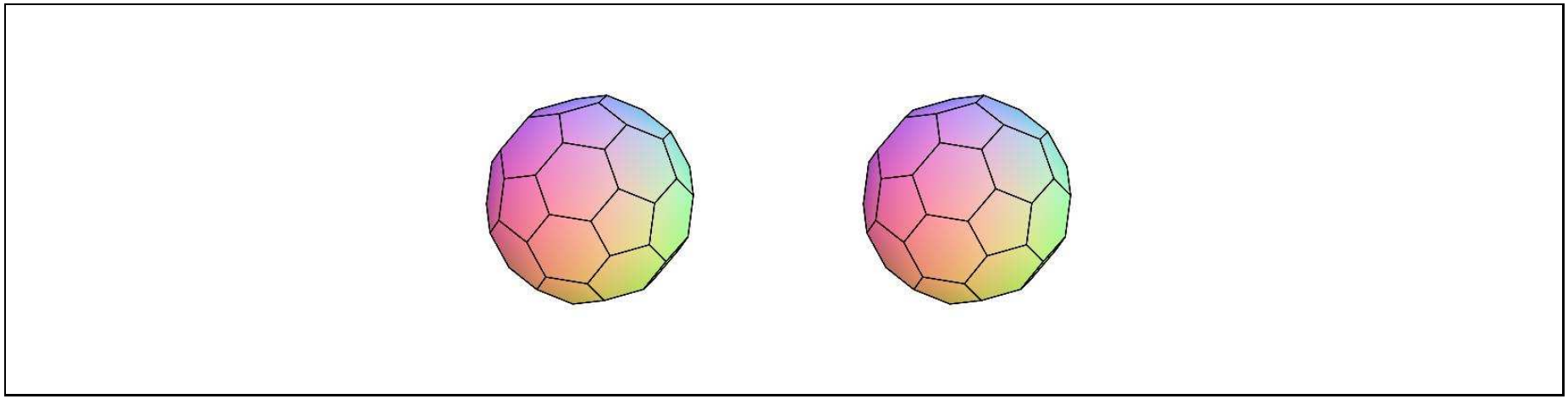


Figure 5.1: This is a symmetry of the soccer ball. We can not tell if this soccer ball been moved.

[I] The student's action is called *symmetry* of the soccer ball. More generally, for some set $X \subseteq \mathbb{R}^3$ such an action is called a symmetry of X. The set of symmetries of X is denoted $S(X)$. Intuitively, the set of such symmetries forms a group under composition: The set of symmetries is closed since if I turned away twice and two students reoriented the soccer ball in succession, I would not be able to tell the ball had been touched. Composition of symmetries are associate since they correspond to the composition of functions from the set which is the soccer ball to itself. There is an identity since if some student came up to the soccer ball and didn't touch it at all, I certainly wouldn't be able to tell it was touched (indeed, in this case I would be correct to guess it was not touched!). Lastly, every symmetry of the soccer ball has an inverse since any symmetry can be undone by doing the original symmetry in reverse. We give a more rigorous proof of this the appendix, Theorem C.29.

For any $X \in \mathbb{R}^3$, $S(X)$ is a group.

REMARK 5.7. The intuition here is roughly:

Roughly, a group is something which preserves some sort of a structure.

In our example, the structure is "distance." If the structure were just "angles" we would also be allowed to dilate objects, via $\mathbf{v} \mapsto 2\mathbf{v}$. This would double the size of objects, but it would not change angles. For instance, consider the difference between "similar" triangles and "congruent" triangles that one learns about in geometry.

It's also easy to identify something that is not a symmetry, because the resulting object "looks different." For instance, see Figure **??**. If a student wanted to be sneaky and move a square on my desk (don't as me why!) while I was not looking and not get caught (I know, the horrors of square moving) they would not move the square according to this picture.

[I] Mathematicians call a soccer ball a truncated icosahedron. Check out wiki's article

http://en.wikipedia.org/wiki/Truncated icosahedron

for more information and some cool graphics.

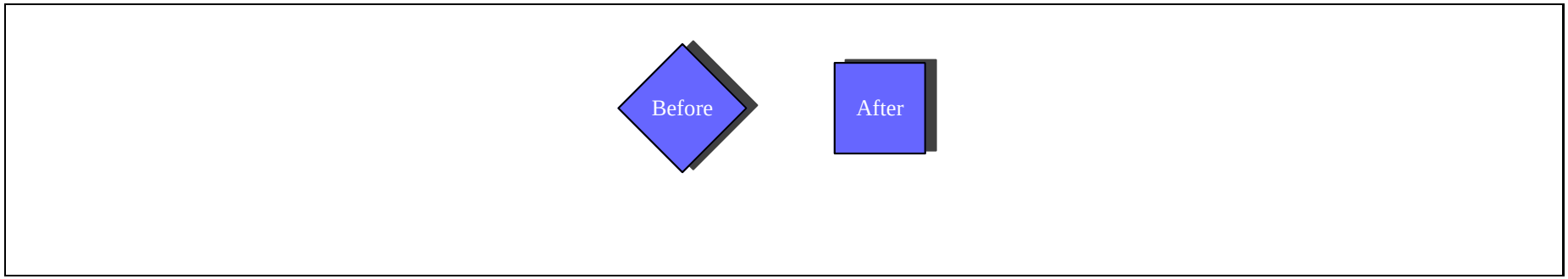


Figure 5.2: This is not a symmetry of the square. We know this square has been moved even if we didn't actually see it get moved.

5.2.1 Dihedral Groups

Let's study a set in more detail to get a better sense of the $S(X)$ groups. Of course taking the soccer ball to be X, while illustrating the point, is too complicated of an example to start out with. In fact there are 60 symmetries of a soccer ball. Let's start instead with a square.[II]

Remark 5.8. Most of the time we will not bother to explain how X sits inside of $\mathbb{R}^3$. For instance, we talk about the symmetries of the square without ever saying how the square sits inside of $\mathbb{R}^2$. It should be intuitive that it really doesn't matter which square we choose inside of $\mathbb{R}^2$. Is it the square with vertices $1, i, -i, -1$ or perhaps the square with vertices $2, 2i, -2i, -2$ or perhaps the square with vertices $1+i, 1-i, -1+i, -1-i$? Perhaps the square is inside of $\mathbb{R}^3$ and not even $\mathbb{R}^2$. Is the red square on the chalk board or the blue square on Mike's graduation hat? It turns out that in a certain sense it doesn't matter how it sits. This should be intuitive if one actually picks up a physical square and plays with it. It doesn't matter if you're in Kansas or the moon, the symmetries of this object will be the "same". [III]

Let's study the group of symmetries of a square, ◆. This group will typically be denoted as $\mathcal{D}_4$ instead of $S(◆)$ (the 4 refers to the number of vertices in a square). There are 8 such symmetries, four of which are rotations and the other four of which are reflections (see below). To demonstrate these 8 symmetries, we label the vertices of the square by the numbers 1, 2, 3, and 4. The picture shown depicts what happens to the square after the symmetry has been applied.
(Note, we can achieve all of the reflections by rotating by 180 degrees along the x axis. For instance r_1 is $r_1(◆) = \varrho_{\mathbf{i},\pi}(◆)$.) For instance, the first picture indicates the identity (i.e. rotation by 0 degrees) the second indicates rotation by 90 degrees in a counter-clockwise manner. Thus, in symbols, the elements of this group are

$$\{\mathbf{1}_◆, \varrho_{\frac{\pi}{2}}, \varrho_{\pi}, \varrho_{\frac{3\pi}{2}}, r_1, r_2, r_3, r_4\}.$$

Now we could make a table out of these guys, but the table wouldn't reveal too much. To make things easier to read, we first derive some basic identities in this group.

Fact 1 $\varrho_{\frac{\pi}{2}}^4 = \mathbf{1} = r_1^2$,

Fact 2 $r_1 \circ \varrho_{\frac{\pi}{2}} \circ (r_1)^{-1} = \varrho_{\frac{\pi}{2}}^{-1} = \varrho_{\frac{3\pi}{2}}$,[IV]

Fact 3 All symmetries can be written in terms of (i.e. as products of) $\varrho_{\frac{\pi}{2}}$ and r_1 as indicated by

$$\begin{aligned} \mathbf{1}_◆ &= \varrho_{\frac{\pi}{2}}^4 & \varrho_{\frac{\pi}{2}} &= \varrho_{\frac{\pi}{2}} & \varrho_{\pi} &= \varrho_{\frac{\pi}{2}}^2 & \varrho_{\frac{3\pi}{2}} &= \varrho_{\frac{\pi}{2}}^3 \\ r_1 &= r_1 & r_2 &= r_1 \circ \varrho_{\frac{\pi}{2}} & r_3 &= r_1 \circ \varrho_{\frac{\pi}{2}}^2 & r_4 &= r_1 \circ \varrho_{\frac{\pi}{2}}^3. \end{aligned} \tag{1}$$

[II] We could start with a regular triangle instead of a square, but this example is not as interesting since there are symmetries which allow you can permute the three vertices of a regular triangle any way you wish. In this sense the symmetries of a triangle are "the same" as $\mathcal{S}_3$.

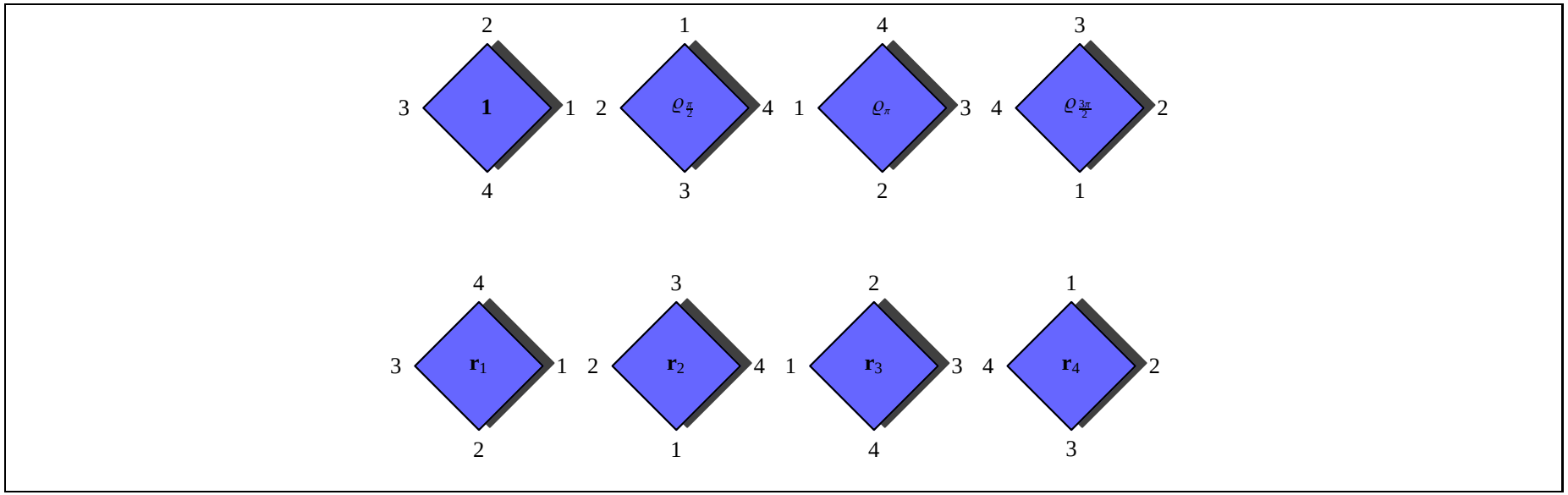


Figure 5.3: The eight symmetries of the square (After the symmetry has been applied.)

Using these facts, we can set the square aside and simply do algebra. For instance,

$$\begin{aligned} r_2 \circ r_3 &= \left(r_1 \circ \varrho_{\frac{\pi}{2}}^3\right) \circ \left(r_1 \circ \varrho_{\frac{\pi}{2}}^2\right) \\ &= r_1 \circ r_1 \circ \varrho_{\frac{\pi}{2}} \circ \varrho_{\frac{\pi}{2}}^2 \\ &= \mathbf{1}_{\blacklozenge} \circ \varrho_{\frac{3\pi}{2}} \\ &= \varrho_{\frac{3\pi}{2}}. \end{aligned}$$

This looks even sexier if we set $\rho_4 = \rho = \varrho_{\frac{\pi}{2}}$ and $\mathbf{r} = r_1$ and only write elements in the form $\mathbf{r}^\epsilon \rho^j$ (with $\epsilon \in \{0,1\}$ and $j \in \{0,1,2,3\}$) Repeating the previous calculation (while omitting all $\circ$ for sanity), this looks like

$$\begin{aligned} \left(\mathbf{r}\rho^3\right)\left(\mathbf{r}\rho^2\right) &= \mathbf{r}\mathbf{r}\rho\rho^2 \\ &= \rho^3. \end{aligned}$$

This approach of using the ρ's and $\mathbf{r}$'s and the basic rules is more algebraic but it is less geometric. With the dictionary established in (1), we can go back and forth whenever we would like. We can now fill the entire multiplication table for this group rather rapidly.

EXAMPLE 5.9. We illustrate how to do some simplifications to even more complicated products. The general rule is that whenever I pass an ρ by a $\mathbf{r}$ I must replace that ρ by an ρ^{-1}. Hence, when I pass a $\mathbf{r}$ by an ρ^i I need to replace it with a

$$\rho^{-i} = \rho^4 \rho^{-i} = \rho^{4-i}.$$

[IV] This is a special case of Corollary C.40 proven in the appendix. For us, we can verify this particular identity by just playing with the square.

(In particular $\mathbf{r}$ and ρ^2 commute).

$$\begin{aligned}
\rho^3\mathbf{r}\rho^2\mathbf{r}\rho\mathbf{r}\rho^3\mathbf{r}\rho &= \rho^3\mathbf{r}\rho^2\mathbf{r}\rho\mathbf{r}\rho^3\rho^3\mathbf{r} \\
&= \rho^3\mathbf{r}\rho^2\mathbf{r}\rho\mathbf{r}\rho^6\mathbf{r} \\
&= \rho^3\mathbf{r}\rho^2\mathbf{r}\rho\mathbf{r}\rho^2\mathbf{r} \\
&= \rho^3\mathbf{r}\rho^2\mathbf{r}\rho\rho^2\mathbf{r}^2 \\
&= \rho^3\mathbf{r}\rho^2\mathbf{r}\rho^3 \\
&= \rho^3\mathbf{r}\rho^2\rho\mathbf{r} \\
&= \rho^3\mathbf{r}\rho^3\mathbf{r} \\
&= \rho^3\rho\mathbf{r}^2 \\
&= \rho^4 \\
&= \mathbf{1}
\end{aligned}$$

and we have simply found a really fancy (and dumb) way of writing the identity!

Here is the entire table for $\mathcal{D}_4$.

$\mathcal{D}_4$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$
$\mathbf{1}$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$
ρ	ρ	ρ^2	ρ^3	$\mathbf{1}$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$
ρ^2	ρ^2	ρ^3	$\mathbf{1}$	ρ	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$
ρ^3	ρ^3	$\mathbf{1}$	ρ	ρ^2	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$
$\mathbf{r}$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{1}$	ρ	ρ^2	ρ^3
$\mathbf{r}\rho$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	ρ^3	$\mathbf{1}$	ρ	ρ^2
$\mathbf{r}\rho^2$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	ρ^2	ρ^3	$\mathbf{1}$	ρ
$\mathbf{r}\rho^3$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	ρ	ρ^2	ρ^3	$\mathbf{1}$

DEFINITION 5.10. The set of symmetries of a regular n-gon is denoted $\mathcal{D}_n$. That is we are defining

$$\mathcal{D}_n \stackrel{\text{def}}{=} S(\{\{n\}\}).$$

The group, $(\mathcal{D}_n, \circ)$ is known as the *dihedral group* of degree n.

For example, we just wrote down the table for $\mathcal{D}_4$ above. For consistency, we make the following definitions, which ought to appear fairly odd at this point.

Convention 5.11. We define $\mathcal{D}_2$ as the symmetries of a rectangle, which is not a square. And we define $\mathcal{D}_1$ as the symmetries of a the line segment $[-1, +1]$.

We will look more closely at $\mathcal{D}_9$ more closely later in the chapter (see page 302). For instance, we will prove that $|\mathcal{D}_n| = 2n$.

Remark 5.12. You should be asking yourselves "why I am not allowing reflections in my definition of $\mathcal{D}_n$?" The reason is quite simple and has already been alluded to. For a two dimensional object, like any polygon, you can reflect it by performing a 180° rotation in three dimensions (along an axis which is in the $x - y$ plane). Indeed, this is the only way that we can physically "reflect" anything! If I asked you to reflect a piece of paper sitting in front of you, how would you do it? You would pick it up, and spin it around and then put it back down. You never actually "reflected", you rotated in one higher dimension! This fits in with the pop mathematical knowledge that if we lived in four dimensions, you could, at least in theory, turn yourself inside out.

5.3 Basic Properties of Groups

The first property that makes a group cooler than a monoid is cancelation.

Theorem 5.13. *Let G be a group and $g, g_1, g_2 \in G$. Then we have*

- **(Left Cancelation)** $g \star g_1 = g \star g_2 \Rightarrow g_1 = g_2$.
- **(Right Cancelation)** $g_1 \star g = g_2 \star g \Rightarrow g_1 = g_2$.

Proof. We have that

$$\begin{aligned} & g \star g_1 = g \star g_2 \\ \Rightarrow & g^{-1} \star (g \star g_1) = g^{-1} \star (g \star g_2) \\ \Rightarrow & (g^{-1} \star g) \star g_1 = (g^{-1} \star g) \star g_2 \\ \Rightarrow & e_G \star g_1 = e_G \star g_2 \\ \Rightarrow & g_1 = g_2. \end{aligned}$$

This proves left cancelation. Right cancelation is proven in the same way, except we now multiply both sides on the right by g^{-1}. ∎

Exercise 5.7. Prove that cross cancelation does not hold in a group. That is show that the statement:

$$g \star h_1 = h_2 \star g$$

does not necessarily imply that $h_1 = h_2$ by finding an explicit counter example.

Solution

Exercise 5.8. Prove that in fact cross cancelation (defined in the previous exercise) holds for all elements of a group iff the group is abelian.

SOLUTION

Remark 5.14.

(1) You should think of the the g's as literally canceling out in the equations $g \star g_1 = g \star g_2$ and $g_1 \star g = g_2 \star g$. In other words

$$g \star g_1 = g \star g_2 \Rightarrow \cancel{g} \star g_1 = \cancel{g} \star g_2 \Rightarrow g_1 = g_2.$$

(2) A fancy way of saying this is by saying for any element g in a group, the functions

$$L_g : h \mapsto g \star h \quad \text{and} \quad R_g : h \mapsto h \star g$$

are injective. This can easily be seen directly since L_g is easily seen to have an inverse (in the function sense), namely $L_{g^{-1}}$ ... to undo left multiplication by g we simply need to left multiply by g^{-1}. This logic also shows that L_g and R_g are also onto. That is for any $g, y \in G$ there exists (by surjectivity) a unique (by injectivity) $x \in G$ which solves the equation

$$g \star x = y.$$

Simply take $x = g^{-1} \star y$.

(3) This is something quite special about groups as opposed to other systems. For instance we have that for any two real numbers r, s (or complex numbers, or integers, or ... etc) $r \cdot 0 = 0 = s \cdot 0$. For a less trivial looking example consider the equality $2 \cdot 2 = 7 \cdot 2$ which occurs in $\mathbb{Z}_2$ (they both equal 4 there).

Corollary 5.15. *If g is an element of a group with $g^2 = g$, then $g = e_G$.*

Proof. Suppose that $g \star g = g$. Then,

$$g \star g = g = g \star e_G \Rightarrow \cancel{g} \star g = \cancel{g} \star e_G \Rightarrow g = e_G.$$ ■

Corollary 5.16 (Sudoku). *In the multiplication table of a finite group $(G, \star)$, every element of G appears exactly once in every row and column.*

Proof. Suppose we have the same entry twice in, say, the i^{th} row in say the j and k columns. Then $g_i \star g_j = g_i \star g_k$ which implies (by left cancelation) that $g_j = g_k$, or $j = k$. Then a pidgin hole argument tells us that each entry must occur in the i^{th} row *at least* once. The statement about columns is proved instead by using right cancelation. ■

This also says that the multiplication table of a group a latin square.

Example /Theorem 5.17. There is only one table possible for a group of order 3.

Proof. This means we are trying to prove that if $(\{e, a, b\}, \star)$ is a group of order 3 (with e the identity) then the multiplication must look a certain way.

For starters, since e is the identity we know that our table must look as follows:

$\star$	e	a	b
e	e	a	b
a	a	$?_1$	$?_3$
b	b	$?_2$	$?_4$

where the question marks still need to be filled in. Now $?_1$ by Sudoku, $?_1$ cannot be a (since there is an a entry directly above the $?_1$ entry) so it has to be one of the other two elements: either e or b. We show that it has to be b by proving that it cannot be e. For if it were e Sudoku says that the entry $?_3$ directly to the right of the e cannot be anything, indeed

- It cannot be a since a lives in the same row (two entries to the left)
- It cannot be e since we are assuming that $?_1 = a$ and so a already shows up in this row (one entry to the left)
- It cannot be b since the entry right above $?_3$ is already b.

This contradiction means that our assumption that $?_1 = e$ was wrong and so we are stuck with the only other possibility, i.e. $?_1 = b$.

$\star$	e	a	b
e	e	a	b
a	a	b	$?_3$
b	b	$?_2$	$?_4$

Everything else now quickly falls into place. The only entry which hasn't been used in the row which starts with an a is e, so by Sudoku we must have that $?_3 = e$.

$\star$	e	a	b
e	e	a	b
a	a	b	e
b	b	$?_2$	$?_4$

The only thing that hasn't been used in the column which starts with a is (also) e. Thus $?_2 = e$.

$\star$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	$?_4$

Lastly, the only thing that hasn't been used in the row which starts with b is a, thus $?_4 = a$.[I] Hence the table must be

$\star$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

.

Among other things, this proves that every group of order 3 is abelian. Notice, if we replace all occurrences of e by $[0]_3$, a by $[1]_3$, and b by $[2]_3$ this table matches exactly with the table for $(\mathbb{Z}_3, +)$

$+$	$[0]_3$	$[1]_3$	$[2]_3$
$[0]_3$	$[0]_3$	$[1]_3$	$[2]_3$
$[1]_3$	$[1]_3$	$[2]_3$	$[0]_3$
$[2]_3$	$[2]_3$	$[0]_3$	$[1]_3$

.

(which it would have to since $(\mathbb{Z}_3, +)$ is a group of order 3, and we just showed there's only one possible!) Thus every table of a group of order 3, in some sense, is equivalent to $\mathbb{Z}_3$. (We will later say that this implies that every group of order 3 is *isomorphic* to $\mathbb{Z}_3$.) ■

EXERCISE 5.9. Suppose that

$\star$	a	b	c	d	e
a	b	?	?	?	d
b	e	d	?	?	c
c	?	?	c	?	?
d	?	?	?	?	?
e	d	?	?	b	a

.

is a partially filled table for a group. Fill in the rest of the table.[II] Which familiar group has a "matching" table (you might have to permute the order the table is usually written)? Which element is the identity element?

Hint! It's not e.

SOLUTION

[I] Alternatively we could have noted that the only thing which hasn't been used in the column which starts with b is (also) a.

Exercise 5.10. There are algebraic systems which satisfy every condition except for associativity and even satisfy Sudoku. For instance, consider

Solution

$\star$	e	a	b	c	d	x	y	z
e	e	a	b	c	d	x	y	z
a	a	e	c	b	x	y	z	d
b	b	c	e	a	y	z	d	x
c	c	b	a	e	z	d	x	y
d	d	z	y	x	e	a	b	c
x	x	d	z	y	c	e	a	b
y	y	x	d	z	b	c	e	a
z	z	y	x	d	a	b	c	e

and show that it has an identity, satisfies Sudoku (just observe that each entry shows up once in every row and column), and has inverses, but show that it is not a group by showing that it is non-associative.

We have already proven the below result for monoids (which all groups are examples of) in general in Theorem 4.34. We give a different proof using cancelation.

Corollary 5.18 (Uniqueness of Inverses=Theorem 4.34)**.** *Let G be a group. Suppose that $h_1, h_2, g \in G$ and that h_1 and h_2 are inverses for g. Then $h_1 = h_2$.*

Proof. The hypothesis guarantees that $g \star h_1 = e_G$ and $g \star h_2 = e_G$ as well. In particular, $g \star h_1 = g \star h_2$ and so

$$\cancel{g} \star h_1 = \cancel{g} \star h_2 \Rightarrow h_1 = h_2$$

as desired. ∎

Corollary 5.19. *If $a \star b = e_G$, then $a = b^{-1}$.*[III]

Proof. We have

$$a \star b = e_G = b^{-1} \star b \Rightarrow a \star \cancel{b} = b^{-1} \star \cancel{b} \Rightarrow a = b^{-1}$$

as desired. ∎

In words, in a group, the above says that if h is a right inverse for g, then it is automatically an (two sided) inverse of g.

5.4 Direct Products

Mathematicians love to find ways of making new things out of old. In group theory, given two groups, the easiest way of making a new group is by the so-called direct product construction.

[III] Of course this isn't by definition since the definition also requires that $b \star a = e_G$.

Definition 5.20. Given two groups $(H, *)$ and $(K, \star)$ we can form a new group called the direct product of $(H, *)$ and $(K, \star)$ as

$$\big(H \times K, (*, \star)\big).$$

In more detail, we define a binary operation, $\bullet \stackrel{\text{def}}{=} (*, \star)$, on the set $H \times K$ via

$$(h, k) \bullet (h', k') \stackrel{\text{def}}{=} (h * h', k \star k').^{\text{I}}$$

Exercise 5.11. In this problem we show that $H \times K$ is a group whenever H and K are. (At each stage we are only assuming what's stated about H and K.)

Solution

(1) Show that with this operation, $(H \times K, \bullet)$ is an algebraic system whenever H and K are.

(2) Now suppose that $(H, *)$ and $(K, \star)$ are semi-groups. Show that for all $h, h', h'' \in H$ and $k, k', k'' \in K$ that

$$((h, k) \bullet (h', k')) \bullet (h'', k'') = (h, k) \bullet ((h', k') \bullet (h'', k''))$$

by multiplying out both sides and using the associativity of $(H, *)$ and $(K, \star)$. Conclude that $(H \times K, \bullet)$ is also a semi-group.

(3) Suppose that $(H, *)$ and $(K, \star)$ have identities e_H and e_K respectively. Prove that (e_H, e_K) is an identity for $(H \times K, \bullet)$. Conclude that if $(H, *)$ and $(K, \star)$ are monoids, then so is $(H \times K, \bullet)$.

(4) Show that $(h, k) \in \mathcal{U}(H \times K, \bullet)$ iff $h \in \mathcal{U}(H, *)$ and $k \in \mathcal{U}(K, \star)$. Conclude that

$$\mathcal{U}(H \times K, \bullet) = \mathcal{U}(H, *) \times \mathcal{U}(K, \star).$$

(5) Conclude that $(H \times K, \bullet)$ is a group whenever both $(H, *)$ and $(K, \star)$ are groups.

Exercise 5.12. Show that $(H \times K, \bullet)$ is abelian iff both $(H, *)$ and $(K, \star)$ are abelian.

Solution

Notation 5.21. We have natural copies of both H and K inside of the set $H \times K$ given by

$$\widetilde{H} \stackrel{\text{def}}{=} H \times \{e_K\} \quad \text{and} \quad \widetilde{K} \stackrel{\text{def}}{=} \{e_H\} \times K.$$

[I] In general, if $f\colon X_1 \to Y_1$ and $g\colon X_2 \to Y_2$ are functions, then

$$(f, g)\colon X_1 \times X_2 \to Y_1 \times Y_2$$

given by $(f, g)(x, x') = (f(x), g(x'))$ defines a function. In our special case $X_1, Y_1 = H$, and $X_2, Y_2 = K$.

That is $\widetilde{H}$ is the set of elements of $H \times K$ whose second component is e_K, i.e.

$$\widetilde{H} = \{(h, e_K) : h \in H\}.$$

A similar statement can be made about $\widetilde{K}$.[II] Occasionally, we are sloppy about this, and we consider H and K to literally be subsets $\widetilde{H}$ and $\widetilde{K}$, respectively. This is something that we should be familiar with even in the case of $\mathbb{R}^n$. For instance, we typically consider $\mathbb{R}$ as a subset of $\mathbb{R}^2$ by thinking of it as the x-axis.

Exercise 5.13. In $H \times K$, prove that every element of $\widetilde{H}$ commutes with every element of $\widetilde{K}$.

Solution

Exercise 5.14. Make a table for $(\mathbb{Z}_3 \times \mathbb{Z}_2, (+,+))$.

Solution

Exercise 5.15. Make a table for $(\mathcal{D}_3 \times \mathbb{Z}_2, (\circ,+))$.

Solution

Example 5.22. The following table is the table for $\mathbb{Z}_2 \times \mathbb{Z}_2$.

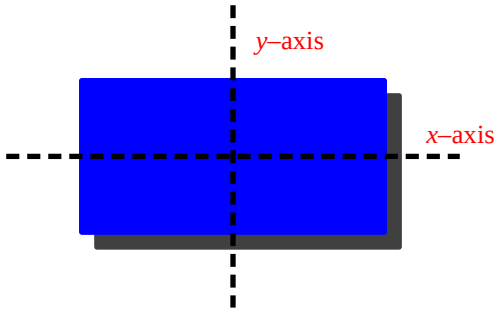
$(+,+)$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$
$([0]_2,[0]_2)$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$
$([1]_2,[0]_2)$	$([1]_2,[0]_2)$	$([0]_2,[0]_2)$	$([1]_2,[1]_2)$	$([0]_2,[1]_2)$
$([0]_2,[1]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$
$([1]_2,[1]_2)$	$([1]_2,[1]_2)$	$([0]_2,[1]_2)$	$([1]_2,[0]_2)$	$([0]_2,[0]_2)$

5.5 More Symmetry Groups

Every object in the universe has symmetry (all be it many have trivial symmetry!). And the set of symmetries of any object forms a group. As such, we have a vast amount of examples of groups at our finger tips.

Example 5.23. Denote a rectangle which is not a square by the symbol ▬. For instance, we could take ▬ to have vertices $(2,1)$, $(-2,1)$, $(-2,-1)$, and $(2,-1)$. Recall, by definition, the symmetries of the rectangle $S(\text{▬})$ (this will also be denoted $\mathcal{D}_2$, the dihedral group of degree 2.)

[II] Of course things like $\widetilde{H}$ and $\widetilde{K}$ don't make any sense outside of the context we are discussing now! Namely direct products!



Now reflection across the two axis are symmetries of this rectangle denoted by r_1 and r_2 respectively. However, the rectangle is less symmetric than the square, and we have lost the symmetries across the diagonals which connect opposite vertices, see Figure 5.4. Such reflections would **not** preserve the rectangle.

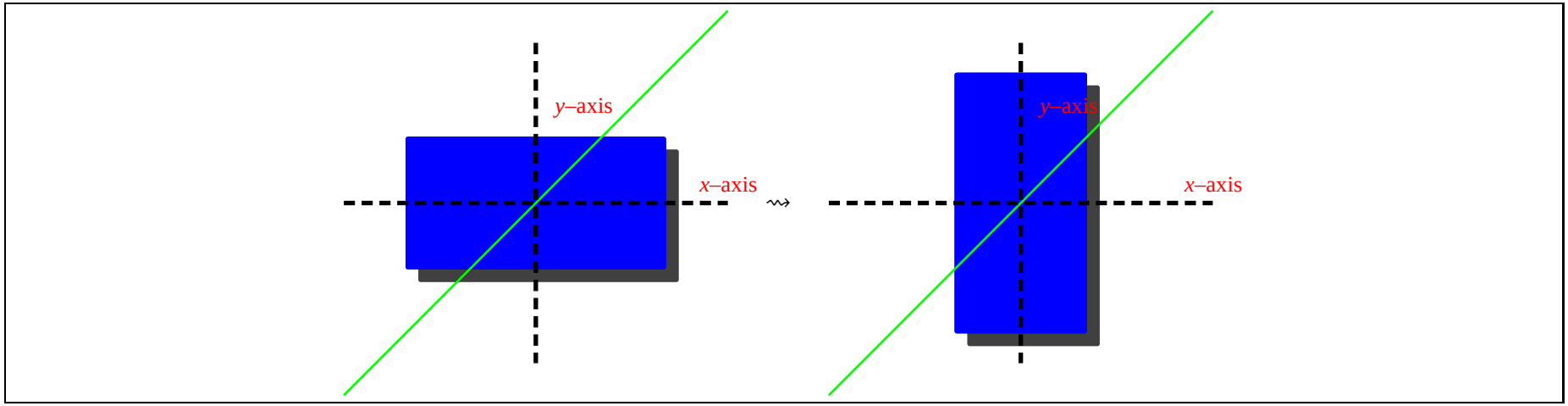


Figure 5.4: This rectangle has been touched.

The last symmetry that we get is the composition of the two listed reflections, which is a rotation by 180 degrees, ϱ_π. This gives us the following multiplication table (the operation is composition of symmetries [which are functions!])

$\circ$	$\mathbf{1}$	r_1	r_2	ϱ_π
$\mathbf{1}$	$\mathbf{1}$	r_1	r_2	ϱ_π
r_1	r_1	$\mathbf{1}$	ϱ_π	r_2
r_2	r_2	ϱ_π	$\mathbf{1}$	r_1
ϱ_π	ϱ_π	r_2	r_1	$\mathbf{1}$

.

Let's compare this table with the table for $(\mathbb{Z}_2 \times \mathbb{Z}_2, (+,+))$ we discussed earlier. It's interesting to note that if we make the following replacements

$$\begin{aligned}([0]_2, [0]_2) &\mapsto \mathbf{1}\\ ([1]_2, [0]_2) &\mapsto r_1\\ ([0]_2, [1]_2) &\mapsto r_2\\ ([1]_2, [1]_2) &\mapsto \varrho_\pi\end{aligned}$$

everywhere they show up in the table for $\mathbb{Z}_2 \times \mathbb{Z}_2$ we end up with this same table!

$(+,+)$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$
$([0]_2,[0]_2)$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$
$([1]_2,[0]_2)$	$([1]_2,[0]_2)$	$([0]_2,[0]_2)$	$([1]_2,[1]_2)$	$([0]_2,[1]_2)$
$([0]_2,[1]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$
$([1]_2,[1]_2)$	$([1]_2,[1]_2)$	$([0]_2,[1]_2)$	$([1]_2,[0]_2)$	$([0]_2,[0]_2)$

$\rightsquigarrow$

$\circ$	$\mathbf{1}$	r_1	r_2	ϱ_π
$\mathbf{1}$	$\mathbf{1}$	r_1	r_2	ϱ_π
r_1	r_1	$\mathbf{1}$	ϱ_π	r_2
r_2	r_2	ϱ_π	$\mathbf{1}$	r_1
ϱ_π	ϱ_π	r_2	r_1	$\mathbf{1}$

.

DEFINITION 5.24. The *Klein 4 group*, denoted V, is the set of permutations

$$\begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix} \quad \begin{pmatrix}1&2&3&4\\3&4&1&2\end{pmatrix} \quad \begin{pmatrix}1&2&3&4\\4&3&2&1\end{pmatrix}$$

and the identity permutation.[I]

EXERCISE 5.16.

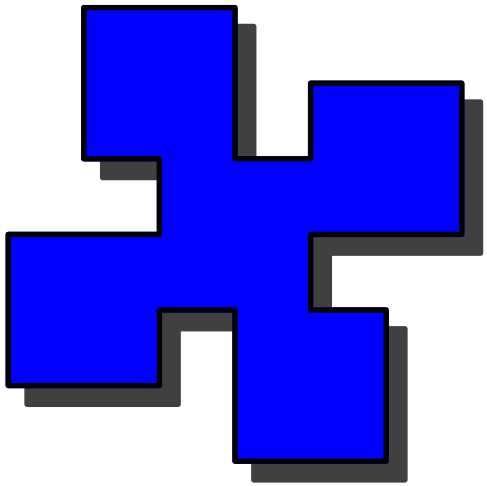
SOLUTION

(1) Write down a multiplication table for the Klein 4 group, V, and verify that it is a group under function composition. Do this by verifying the axioms.

(2) Show that the table for V "matches" the tables for $\mathbb{Z}_2^2$ and $S(\blacksquare)$ in the sense above.

EXAMPLE 5.25. Next consider the shape which appears below.

[I] The letter V stands for *Vierergruppe*, which is german for "group of four."



Here, all symmetries of this object are direct – i.e., there are no reflections which preserve the shape, however there are four rotations (as usual we're counting the identity as a rotation by angle 0),

$$\mathbf{1}, \varrho_{\pi/2}, \varrho_{\pi}, \varrho_{3\pi/2}.$$

Together they form the following table:

$\circ$	$\mathbf{1}$	$\varrho_{\frac{\pi}{2}}$	ϱ_{π}	$\varrho_{\frac{3\pi}{2}}$
$\mathbf{1}$	$\mathbf{1}$	$\varrho_{\frac{\pi}{2}}$	ϱ_{π}	$\varrho_{\frac{3\pi}{2}}$
$\varrho_{\frac{\pi}{2}}$	$\varrho_{\frac{\pi}{2}}$	ϱ_{π}	$\varrho_{\frac{3\pi}{2}}$	$\mathbf{1}$
ϱ_{π}	ϱ_{π}	$\varrho_{\frac{3\pi}{2}}$	$\mathbf{1}$	$\varrho_{\frac{\pi}{2}}$
$\varrho_{\frac{3\pi}{2}}$	$\varrho_{\frac{3\pi}{2}}$	$\mathbf{1}$	$\varrho_{\frac{\pi}{2}}$	ϱ_{π}

.

It's interesting to note if we take the table for $\mathbb{Z}_4$ and make the following replacements

$$\begin{aligned}
[\,0\,]_4 &\mapsto \mathbf{1} \\
[\,1\,]_4 &\mapsto \varrho_{\pi/2} \\
[\,2\,]_4 &\mapsto \varrho_{\pi} \\
[\,3\,]_4 &\mapsto \varrho_{3\pi/2}
\end{aligned}$$

everywhere they show up in the table for $\mathbb{Z}_4$ we end up with this same table!

$+$	$[0]_4$	$[1]_4$	$[2]_4$	$[3]_4$
$[0]_4$	$[0]_4$	$[1]_4$	$[2]_4$	$[3]_4$
$[1]_4$	$[1]_4$	$[2]_4$	$[3]_4$	$[0]_4$
$[2]_4$	$[2]_4$	$[3]_4$	$[0]_4$	$[1]_4$
$[3]_4$	$[3]_4$	$[0]_4$	$[1]_4$	$[2]_4$

$\rightsquigarrow$

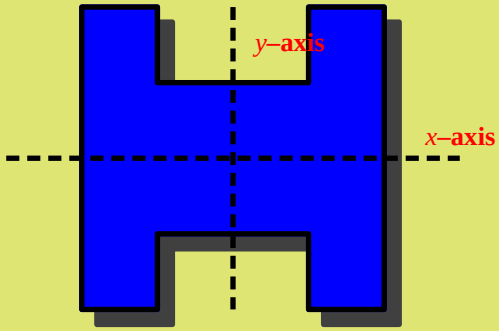
$\circ$	$\mathbf{1}$	$\varrho_{\frac{\pi}{2}}$	ϱ_{π}	$\varrho_{\frac{3\pi}{2}}$
$\mathbf{1}$	$\mathbf{1}$	$\varrho_{\frac{\pi}{2}}$	ϱ_{π}	$\varrho_{\frac{3\pi}{2}}$
$\varrho_{\frac{\pi}{2}}$	$\varrho_{\frac{\pi}{2}}$	ϱ_{π}	$\varrho_{\frac{3\pi}{2}}$	$\mathbf{1}$
ϱ_{π}	ϱ_{π}	$\varrho_{\frac{3\pi}{2}}$	$\mathbf{1}$	$\varrho_{\frac{\pi}{2}}$
$\varrho_{\frac{3\pi}{2}}$	$\varrho_{\frac{3\pi}{2}}$	$\mathbf{1}$	$\varrho_{\frac{\pi}{2}}$	ϱ_{π}

.

This really shouldn't be too terribly shocking since you can "think" of addition in $\mathbb{Z}_4$ as adding 90° rotations. What might be more surprising is that there is no way to replace entries of the table of the last group we consider, the symmetries of a (proper) rectangle and get either of these tables. In a sense the difference between these two tables is that in the first two tables we had the identity going straight down the diagonal and in this second pair of tables we do not.

EXERCISE 5.17. In the imprecise way just discussed, which table is the symmetries of the letter *H* "the same" as? Why?

SOLUTION



5.6 TYPICAL SHORTHANDS

Typically, for a group $(G, \star)$, the binary operation $\star$ is inferred, and we simply refer to G as group. If we did this when we were studying monoids in the previous chapter this would have been a disaster! For instance, the set of functions $\mathcal{F}_{\mathbb{R}} = \{f \mid f : \mathbb{R} \to \mathbb{R}\}$ can be made into a monoid under three different operations, addition of functions, multiplications of functions, and composition of functions. More practically, all the rings we've studied have two binary operations which make them into monoids. However, in the realm of group theory there seems to be very few concrete examples, and none that we shall study, of a case where the same set has two naturally arising binary operators which form a group. For instance the sets $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{H}$, $\mathbb{Z}_n$, and $\mathbb{C}$ are only groups under addition (and not multiplication) since all of these sets contain 0. (And 0 cannot have a multiplicative inverse for otherwise $1 = 0 \cdot 1 = 0$ implies that $1 = 0$.) Thus, for instance, when we talk about the group $\mathbb{Z}_n$ we **really** mean the group $(\mathbb{Z}_n, +)$. On the other hand, if we look at the set of invertible elements in all of these sets (remember in all cases except for $\mathbb{Z}$ and $\mathbb{Z}_n$ this means the set of non-zero elements) these sets do not form a group under addition (since in all these cases 1 and -1 are invertibe, but $1 + (-1) = 0$ which is not invertible) when we refer to the group of invertible elements in any of these sets it is implied that we are talking about multiplication. For instance, the groups $\mathbb{R}^\times$, $\mathbb{C}^\times$, and $\mathbb{H}^\times$ imply that the binary operation is multiplication. Similary, the group $\mathcal{U}(\mathbb{Z}_n)$ of classes of integers relatively prime to n really means the group $(\mathcal{U}(\mathbb{Z}_n), \cdot)$ where $\cdot$ means mod n *multiplication*. If the given set is a set of functions, (as is the case for $\mathcal{S}_n$ and $\mathcal{D}_n$) the implicit binary operation is composition, $\circ$. Thus "the group $\mathcal{S}_n$" really means the group $(\mathcal{S}_n, \circ)$. While this notation is technically sloppy, it will help us in the long run.

To make matters worse (or better if you don't like writing!) we leave off the operation $\star$ even when performing products in an abstract group (unless we happen to have $\star = +$). So for instance we refer to a group $(G, \star)$ as G and the product of two elements a and $b \in G$ as ab instead of $a \star b$.

CONVENTION 5.26. From now on, unless we wish to stress something, we shall refer to the set G itself as a group whenever $(G, \star)$ is a group for some binary operation $\star$. We also write products without referencing $\star$ as well.

5.7 POWERS

Computing orders of elements in a group is going to be one of our most realiable methods for telling the difference between different groups. We start this section by talking about powers of elements in a group. We have already talked about positive powers in a monoid, but a group has the added bonus that every element has an inverse. This makes it possible for us to define negative powers as well as positive powers of an element of a group.

Recall that by generalized associativity, the product

$$\underbrace{g\,g\cdots g}_{n\,\text{times}}$$

makes sense and was Definition 4.21 back on page 222, of g^n. If the group operation is addition (meaning denoted by the symbol $+$), as is the case for the groups $\mathbb{Z}_n$, $\mathbb{Z}_n \times \mathbb{Z}_m$, and $\mathbb{Z}$, note that we are looking at

$$\underbrace{g+g+\cdots+g}_{n\,\text{times}}$$

which, in this very special case, will be denoted as ng (or $n \cdot g$) instead of g^n.

EXERCISE 5.18. Let a and b be integers and let n be a positive integer. Prove that

$$a\,[\,b\,]_n = b\,[\,a\,]_n = [\,a\,]_n\,[\,b\,]_n = [\,ab\,]_n\,.$$

SOLUTION

Careful! It is probably more important (and possibly harder for many students) to realize why they "might" be different than to believe that they are the same. Be careful that you don't take anything for granted here!

For some proofs, it is useful to give an inductive definition of g^n.

- $g^0 \stackrel{\text{def}}{=} e_G$.
- Assuming that g^n (for positive integer n) has been defined, then $g^{n+1} \stackrel{\text{def}}{=} g^n \cdot g$.

Taking this definition to heart, g^n is defined to be

$$\underbrace{(((\cdots(g\,g)g)\cdots)\,g}_{n\text{times}}\,.$$

General associativity says we don't need to give a hoot about parenthesis, so these two definitions coincide. It's worth pointing out that we have, by definition, $g^1 = g$. In a group we can extend the current power notation to include negative powers as well.

For $g \in G$, how do we define something like g^{-5}? Well, every freshman math student knows that 3^{-5} really means

$$\frac{1}{3^5} \quad \text{or} \quad \left(\frac{1}{3}\right)^5 = (\text{The multiplicative inverse of } 3)^5 .$$

We generalize this for an abstract group by defining g^{-5} as $\left(g^{-1}\right)^5$. And more generally, we define

Definition 5.27. If n is a positive number, then

$$g^{-n} \stackrel{\text{def}}{=} \left(g^{-1}\right)^n .$$

Of course this makes sense since for any element of a group g^{-1} is an element of a group, and for any element of a group, any positive power of that element has already been defined. You should also note that up to this very moment, the left hand side of this equation was not defined. This means for m a negative integer $-m$ is a positive integer, and we have defined $g^m \stackrel{\text{def}}{=} (g^{-m})^{-1}$.

Generalized socks and shoes, Exercise 4.15, (applied to the case that all $g_i = g$) tells us that we could have well defined g^{-n} as $(g^n)^{-1}$. In other words by taking powers of g and the inverting the result. In more detail,

$$\begin{aligned}(g^n)^{-1} &= \left(\underbrace{g\,g\cdots g}_{n \text{ times}}\right)^{-1} \\ &= \underbrace{g^{-1}\cdots g^{-1}}_{n \text{ times}} \\ &= \left(g^{-1}\right)^n .\end{aligned}$$

Now all integer powers of g are defined.

Therefore, we have proven

Proposition 5.28. *For any positive integer n we have that $(g^n)^{-1} = \left(g^{-1}\right)^n$ and either one could be used to define g^{-n}.* ■

Corollary 5.29. $g^n = e_G \Leftrightarrow \left(g^{-1}\right)^n = e_G$.

Proof. Since $\left(g^{-1}\right)^{-1} = g$ we only need to prove one direction: "if $g^n = e_G$, then $\left(g^{-1}\right)^n = e_G$" and replace g with g^{-1} to prove the other direction. ■

Remark 5.30. Throughout these notes we will move between the three, $(g^n)^{-1}$, g^{-n} and $\left(g^{-1}\right)^n$ without any further notice.

Lemma 5.31. *For any positive integer n we have that $g^{-(n+1)} = g^{-n}g^{-1}$.*

Proof. We have

$$g^{-(n+1)} = \left(g^{-1}\right)^{n+1} = \left(g^{-1}\right)^n g^{-1}.$$

where the first equality is the definition of negative powers and the second equality is by definition of positive powers. ■

If we are careful enough, the following lemma is straight foward homework assignment:

THEOREM 5.32. *For all integers n and m (not necessarily positive) and all elements a in any group G we have*

(1) $e_G^n = e_G$;

(2) $a^n a^m = a^{n+m}$;

(3) $(a^n)^m = a^{nm}$.

Now assume that $b \in G$ commutes with the original element a, then

(4) *a^n and b^m also commute;*

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

PROOF.

(1) Since $e_G^{-1} = e_G$ we have that for all integers n that $e_G^n = e_G^{-n} = e_G^{|n|}$. Thus

$$\underbrace{e_G\, e_G \cdots e_G}_{|n| \text{ times}} = e_G.$$

(2) Exercise 4.10 back on page 222 from the previous chapter established the result for positive integers n and m. If both n and m happened to be negative, then $-n$ and $-m$ would be positive and we'd have that

$$\begin{aligned} a^n a^m &= \left(a^{-1}\right)^{-n} \left(a^{-1}\right)^{-m} \\ &= \left(a^{-1}\right)^{(-n)+(-m)} \\ &= a^{n+m}. \end{aligned}$$

Therefore we only need to consider the case that one of n is positive and the other is negative. Then, in this case, by possibly replacing a by a^{-1} as above, we can assume that n is positive and m is negative. By replacing m by $-m$ we wish to prove that

$$a^n a^{-m} = a^{n-m}$$

whenever n and m are both positive. We now proceed to prove this by induction on m. The case $m = 1$ is lemma (with $n = n - 1$ and $m = 1$). If we assume the result for all $m < k$, then

$$\begin{aligned} a^n a^{-k} &= a^n a^{-(k-1)} a^{-1} \\ &= a^{n-k+1} a^{-1} \\ &= a^{n-k} \end{aligned}$$

where we've used the lemma to get the first and last equality and the induction hypothesis to get the middle.

(3) Note that if m is positive, then

$$(a^n)^{-m} = \left((a^n)^{-1}\right)^m = \left(\left(a^{-1}\right)^n\right)^m$$

and so by replacing a by a^{-1} we can assume that m is positive and we then can (and will) proceed by induction on m. The case that $m = 1$ is trivial. If the result holds for the case that $m < k + 1$, then

$$\begin{aligned}(a^n)^{k+1} &= (a^n)^k a^n && \text{(Previous part)}\\ &= a^{nk}a^n && \text{(Induction Hypothesis)}\\ &= a^{nk+n} && \text{(Previous part)}\\ &= a^{n(k+1)}\end{aligned}$$

and the induction on m (and hence the proof) is complete.

(4) Exercise (we will prove something a bit more abstract more easily and this will easily imply this later, see page 328.)

(5) We first assume that n is positive and proceed by induction on n. The case that $n = 1$ holds by definition. Now assume that the result holds for all $n < k + 1$. Then

$$\begin{aligned}(ab)^{k+1} &= (ab)^k ab\\ &= a^k b^k ab\\ &= a^k ab^k b && (\textbf{(4)} \Rightarrow a, b^k \text{ commute})\\ &= a^{k+1}b^{k+1}.\end{aligned}$$

Therefore, the case that $n = k + 1$ has been proven, and we are done with the case that n is positive.

On the other hand, if n were negative, then by the previous part, **(4)**, if a and b commute then a^{-1} and b^{-1} also commute. By socks and shoes, (Exercise 4.15 back on page 229) and since this part has already been proven for the positive $-n$ we have

$$\begin{aligned}(ab)^n &= \left((ab)^{-1}\right)^{-n}\\ &= \left(b^{-1}a^{-1}\right)^{-n}\\ &= \left(a^{-1}b^{-1}\right)^{-n}\\ &= a^n b^n.\end{aligned}$$

■

Since a commutes with itself, this lemma implies that any power of a commutes with any other power of a (there are two parts of this lemma which imply this!)

EXERCISE 5.19. Prove part **(4)** of the theorem above.

SOLUTION

In this exercise we show that **(5)** is **always** false when a and b do not commute.

EXERCISE 5.20. Let a and b be elements of a group G. Prove that if $(ab)^2 = a^2 b^2$, then a and b commute.

SOLUTION

Exercise 5.21. Prove for every $(h,k) \in H \times K$ any an integer n that $(h,k)^n = (h^n, k^n)$.

SOLUTION

5.8 Additive and Multiplicative Notations.

We will break the notation for a group in half. This will allow us less opportunity to mess up what something along the lines of what 2^3 means in, say, $\mathbb{Z}_2$. We will always insist that this means 8 and never $2 \cdot 3 = 6$. We will basically have two notations for groups: one when the abstract operation is denoted with a plus sign, $+$, generically called addition. Groups whose operation is addition will be called *additive groups.* And one for all other kinds of groups, i.e. when the operation is denoted as $*$ (as in quaternion multiplication), $\cdot$, $\circ$, generically called multiplication. In these notes we will only use additive notation for an abelian group, but many abelian groups, such as $\mathcal{U}(\mathbb{Z}_n)$, will not use additive notation. For multiplicative notation, we will stick to leaving off the operation when the context is clear. For additive notation we have several important distinctions in notation which we collect in this little table:

Multiplicative Notation	Additive Notation
gh	$g+h$
g^{-1}	$-g$
g^n	ng
$e, e_G,$ or $1, 1_G$	$0, 0_G$

.

For instance, the properties written in multiplicative notation, such as $g^{n+m} = g^n g^m$ takes the form of the distributive property when written in additive notation: $(n+m)g = ng + mg$. And the property that $g^{nm} = (g^n)^m$ resembles the associative property when written in additive notation: $(n \cdot m)g = n(mg)$.

Remark 5.33. This notation still doesn't save us when discussing functions from $\mathbb{R} \to \mathbb{R}$. There, one often mixes what these things mean! For instance, $\sin^2 x$ always refers to the operation which is multiplication, while $\sin^{-1} x$ typically refers to the operation which is composition. There really should be a third notation in this crappy case. Maybe something like $\sin^{\circ 3}(x)$ means $\sin(\sin(\sin(x)))$ while something like $\sin^{\cdot -1}(x)$ means $1/\sin x$. However, of course, since both operations are equally important, people will end up dropping the extra notation and we will still end up with confused calculus students.

Warning! 5.34. At least initially, it's worth constantly asking yourself where the various multiplications are taking place! For instance, in the equality

$$(n \cdot m)g = n(mg)$$

in an additive group we just mentioned, the multiplication $n \cdot m$ is the multiplication of integers while on the right hand side, in a sense, both multiplications take a pair of an integer and an element of a group and product another element of a group (that is it's a function $\mathbb{Z} \times G \to G$). Notice that we never defined, and never will define things along the lines of n^g (integer raised to an element of a group), g^n for an additive group, ng for a multiplicative group, and heaven forbid, things like $\ln g$ ... (I have seen students use properties of natural logs to "prove" things about groups before. Please don't be one of them!) Also, in a generic group there is only **one** operation. Therefore, questions like "for any group, is it true that $g(h_1 + h_2) = gh_1 + gh_2$" make as much sense as asking if every frizzle is a wizzle (which doesn't make any sense at all!). This is an example of a question if asked of the math god, the math god would not be able to answer since the math god can only answer questions in terms of things which are defined.

5.9 Order of Elements of a Group; First Definitions.

Here is one of the most important definitions in the course.

Definition 5.35. Let G be a group and $g \in G$.

(1) The element g is said to have *finite order* if there exists a positive integer n so that $g^n = e_G$. Otherwise, g is said to have *infinite order*.

(2) Suppose that g has finite order. Then the least, positive integer n so that $g^n = e_G$ is called the *order* of g. We denote this by writing $|g| = n$.[I]

Exercise 5.22. The following statement (and hence its proof) is false. Find the error in the proof.

Solution

False Theorem 5.36. *Every element of $\mathbb{Z}_6$ has order 6.*

"**Proof**". Let $[a]_6$ be an element of $\mathbb{Z}_6$. Then

$$6[a]_6 = [6]_6 \cdot [a]_6 = [0]_6 \cdot [a]_6 = [0]_6 = e_{\mathbb{Z}_6}.$$

Therefore, by definition, every element of $\mathbb{Z}_6$ has order 6. ■

Exercise 5.23. Prove that an element of a group has infinite order iff every (non-zero) power of that element also has infinite order.
Hint! It will probably look nicer to consider things of finite order instead.

Solution

Remark 5.37.

(1) By definition the order of an element of a group is positive. Therefore, we know instantly that a statement such as $|g| = 0$ cannot possibly be true since the number 0 is not positive. This is despite the fact that we always have $g^0 = e_G$ for any element of any group.

(2) Since the identity e_G of any group G satisfies $e_G^1 = e_G$, and since 1 is the smallest positive integer, $|e_G| = 1$. Conversely, if $g \in G$ has order 1, then we must have that $g^1 = e_G$. But since $g^1 = g$ this means that $g = e_G$, and so the only element of order 1 in any group is the identity.

$$|g| = 1 \Leftrightarrow g = e_G.$$

(3) In practice, to find the order of an element we compute

$$g \quad g^2 \quad g^3 \ldots$$

until we reach the identity for the first time. The exponent of this product is the order. In there is no such integer, then the element has infinite order. [II]

(4) We therefore have two properties for the order of a group:

- $g^{|g|} = e_G$.
- If $0 < m < |g|$, then $g^m \neq e_G$.

(5) After first showing that $g^n = e_G$, the first line of the second part of a proof which would prove that $|g|$ is a particular integer n, we say "suppose that $g^m = e_G$ for some positive integer m." And the last line of such a proof would say "therefore, $m \leq n$."

(6) The order of an element does not depend on which group the element is in (as long as the operations in the different groups are the same ...). For instance, the order of a matrix in $\mathbf{SL}_2(\mathbb{Z})$ has the same order if we consider it as an element in $\mathbf{SL}_2(\mathbb{Z})$, $\mathbf{SL}_2(\mathbb{Q})$, $\mathbf{SL}_2(\mathbb{R})$, or $\mathbf{SL}_2(\mathbb{C})$. In some situations, however, we do need to pay close attention to what operation we are talking about. The most important case comes about in the two groups $\mathbb{Z}_n$ and $\mathcal{U}(\mathbb{Z}_n)$ where the order of the very same element under one of the operations (addition or multiplication) has nothing to do with the order of that very same element under the other operation. In this case, if we wish to make it perfectly case which operation we are talking about, we use the notations $|[a]_n|_+$, or $|[a]_n|_\cdot$. For instance,

$$|[1]_2|_+ = 2 \text{ but } |[1]_2|_\cdot = 1.$$

THEOREM 5.38. *Let g be an element of a group. Suppose that $n = |g|$. Then*

(1) *$|g| = n$ as well; (i.e. $\left|g\right| = \left|g^{-1}\right|$).*

(2) *$g^m = e_G \Leftrightarrow n \mid m$;*

(3) *$g^a = g^b \Leftrightarrow a \equiv_n b$.*

PROOF.

(1) Since e_G^{-1} is e_G itself, and since inverses are unique, for any element $h \in G$, we have that $h^{-1} = e_G$ holds iff $h = e_G$. Therefore, since $(g^m)^{-1} = g^{-m}$ we have that $g^m = e_G$ holds iff $g^{-m} = e_G$. As such the least positive integer m which satisfies $g^m = e_G$ is also the least positive integer m with $g^{-m} = e_G$, i.e. the order of g is the order of g^{-1}.

(2) Suppose that the statement is false, i.e. that $n \nmid m$. Then the division algorithm allows us to write $m = qn + r$ with $1 \leq r < n$ ($r \neq 0$ since we're assuming that $n \nmid m$). Now since

$$\begin{aligned} e_G &= g^m \\ &= g^{qn+r} \\ &= g^{qn}\, g^r \\ &= (g^n)^q\, g^r \\ &= {e_G}^q\, g^r \\ &= e_G\, g^r \\ &= g^r \end{aligned}$$

we have that $g^r = e_G$. But this is impossible since r is a positive integer less than n. The converse is easier and left to you (actually the proof of the converse is contained in the proof of this!).

[II] Of course unless we come up with some theoretical reason why the element has infinite order we could end up, at least in theory, computing powers of g forever!

(3) Suppose that $g^b = g^a$. Then we have that

$$g^b g^{-a} = g^a g^{-a} \Rightarrow g^{b-a} = e_G.$$

But then the previous part implies that $n \mid (b-a)$. But this is the very definition of the notation $a \equiv_n b$. Again, the converse is easier and left to you. ■

EXERCISE 5.24. Suppose that g is an element of a group G and and that $g^{60} = e_G$. What are the possible orders of g?

SOLUTION

EXERCISE 5.25.

(1) Suppose that $g^9 = e_G$ but $g^3 \neq e_G$. Prove that $|g| = 9$.

(2) Suppose that $g^{36} = e_G$ but $g^6 \neq e_G$. Is it necessarily true that $|g| = 9$? Is it necessarily true that $|g| = 36$?

SOLUTION

The following Corollary makes proving that an element has prime order a bit easier. It says to show that an element g of a group has prime order p you simply need to show that g is **not** the identity and g^p **is** the identity. This will be used quite a bit since it is sometimes a pain to stop and show that all previous powers of an element are not the identity, as we will still need to do when showing that an element has order which is not a prime.

COROLLARY 5.39. *Let g be an element of a group G. Suppose that $g \neq e_G$ and that $g^p = e_G$ for some prime p. Then $|g| = p$.*

PROOF. By the theorem, the order of g divides p. But since p is prime, the only positive divisors of p are 1 and p. The order of g cannot be 1 since $g \neq e_G$ by assumption. Therefore the order of g is p as desired. ■

More generally, this says that if we have that g^n is the identity, then to check n is indeed the order, we simply need to make sure that no positive (proper) divisors of n also have that same property.

EXERCISE 5.26. Suppose that g has finite order and r is some integer. Prove that $|g^r| \Big| |g|$.

SOLUTION

In words, this exercise guarantees that the order of any power (or multiple in additive notation) of an element of a group has order less than or equal to the original element's.

DEFINITION 5.40. Let n be an integer. Define

(1) $G(n) \stackrel{\text{def}}{=} \{g \in G : |g| = n\}$ and $G[n] \stackrel{\text{def}}{=} \{g \in G : g^n = e_G\}$.

(2) $\mathbf{o}_n(G) \stackrel{\text{def}}{=} \#(G(n))$ and $\mathbf{O}_n(G) \stackrel{\text{def}}{=} \#(G[n])$

Note that by definition, if $|g| = n$, then $g^n = e_G$, so $G(n) \subseteq G[n]$ and we always have that $\mathbf{o}_n(G) \leq \mathbf{O}_n(G)$. In fact, for all integers > 0, we know that $e_G \in G[n]$ but $e_G \notin G(n)$, therefore for such integers we have strict inequality $\mathbf{o}_n(G) < \mathbf{O}_n(G)$.

Theorem 5.38 can be interpreted in several ways. In particular, Theorem 5.38 **(2)** says that

$$G[n] = \bigcup_{d|n} G(d). \tag{2}$$

Moreover notice that since an element of a group can have only one order ($g \mapsto |g|$ is a function), we have that the $G(n)$'s partition G, i.e.

$$n = m \Leftrightarrow G(n) \cap G(m) \neq \emptyset. \tag{3}$$

Therefore the union in (2) is disjoint and so we have that

$$\mathbf{O}_n(G) = \sum_{d|n} \mathbf{o}_d(G).$$

EXERCISE 5.27. Prove that

$$G[n] \cap G[m] = G[\gcd(n, m)].$$

SOLUTION

5.9.1 LOTS OF EXAMPLES

EXAMPLE 5.41.

(1) $[1]_n \in \mathbb{Z}_n$ has order n since

- $n \cdot [1]_n = [n]_n = [0]_n$.
- Now suppose that m is a positive integer with $[m]_n = m \cdot [1]_n = 0$. Then we have that $n \mid m$. But since $m > 0$ this implies that $m \geq n$.

(2) The rotation by an angle $\frac{2\pi}{n}$, (i.e. the rotation $\varrho_{\frac{2\pi}{n}}$) has order n.

(3) Any reflection has order 2. (Note that rotation by 180° also has order 2, but it is not a reflection.)

(4) Let σ be the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}.$$

Then σ has order 4 since

$$\sigma^2(1) = (\sigma \circ \sigma)(1) = \sigma(2) = 3 \qquad \sigma^2(2) = (\sigma \circ \sigma)(1) = \sigma(3) = 4$$
$$\sigma^2(3) = (\sigma \circ \sigma)(3) = \sigma(4) = 1 \qquad \sigma^2(4) = (\sigma \circ \sigma)(4) = \sigma(1) = 2$$

which doesn't equal the identity, and

$$\sigma^3(1) = (\sigma \circ \sigma \circ \sigma)(1) = \sigma(3) = 4 \qquad \sigma^3(2) = (\sigma \circ \sigma \circ \sigma)(1) = \sigma(4) = 1$$
$$\sigma^3(3) = (\sigma \circ \sigma \circ \sigma)(3) = \sigma(1) = 2 \qquad \sigma^3(4) = (\sigma \circ \sigma \circ \sigma)(4) = \sigma(2) = 3$$

which also does not equal the identity, but

$$\sigma^4(1) = (\sigma \circ \sigma \circ \sigma \circ \sigma)(1) = \sigma(4) = 1 \qquad \sigma^4(2) = (\sigma \circ \sigma \circ \sigma \circ \sigma)(1) = \sigma(1) = 2$$
$$\sigma^4(3) = (\sigma \circ \sigma \circ \sigma \circ \sigma)(3) = \sigma(2) = 3 \qquad \sigma^4(4) = (\sigma \circ \sigma \circ \sigma \circ \sigma)(4) = \sigma(3) = 4$$

which does equal the identity (since it sends every element of $\overline{4}$ to itself).

Hence $\sigma^4 = \mathbf{1}_{\overline{4}}$ but $\sigma, \sigma^2, \sigma^3$ do not. Thus 4 is the least positive integer with that property and by definition the order of σ is 4, i.e. $|\sigma| = 4$. (We could have simply checked that $\sigma^2 \neq \mathbf{1}_{\overline{4}}$ (since, for example, it can quickly be eye–balled to see that it does not fix the letter 1) and that $\sigma^4 = \sigma^2 \circ \sigma^2$ does. This would have cut down a calculation.) Later, we will learn a way to be able to tell what the order of a permutation is on sight (assuming we can calculate least common multiples quickly).

(5) Consider $[\,3\,]_6 \in \mathbb{Z}_6$. then

$$1 \cdot [\,3\,]_6 = [\,3\,]_6 \qquad 2 \cdot [\,3\,]_6 = [\,0\,]_6\,.$$

Hence $|\,[\,3\,]_6\,| = 2$.

(6) Consider $[\,3\,]_{13} \in \mathbb{Z}_{13}$. Then (leaving off square brackets to preserves my sanity)

$$1 \cdot 3 = 3 \qquad 2 \cdot 3 = 6 \qquad 3 \cdot 3 = 9$$
$$4 \cdot 3 = 12 \qquad 5 \cdot 3 = 2 \qquad 6 \cdot 3 = 5$$
$$7 \cdot 3 = 8 \qquad 8 \cdot 3 = 11 \qquad 9 \cdot 3 = 1$$
$$10 \cdot 3 = 4 \qquad 11 \cdot 3 = 7 \qquad 12 \cdot 3 = 10$$
$$13 \cdot 3 = 0.$$

Hence

$$|\,[\,3\,]_{13}\,| = |\,[\,3\,]_{13}\,|_+ = 13.$$

(7) Again consider $[\,3\,]_{13}$ but this time we look for its order as an element of (the multiplicative group!) $\mathcal{U}(\mathbb{Z}_{13})$. Since the idenity of *this* group is $[\,1\,]_{13}$ we are looking for the smallest positive integer n with $([\,3\,]_{13})^n = [\,1\,]_{13}$. We have

$$([\,3\,]_{13})^1 = [\,3\,]_{13} \qquad ([\,3\,]_{13})^2 = [\,9\,]_{13} \qquad ([\,3\,]_{13})^3 = [\,27\,]_{13} = [\,1\,]_{13}\,.$$

Hence

$$|\,[\,3\,]_{13}\,| = |\,[\,3\,]_{13}\,|_\cdot = 3.$$

(The difference between this result and the result from the previous part, Part 6, reiterates that when talking about order of elements in $\mathbb{Z}_n$ we must pay attention to both the n and the operation!)

We could have done Part 6 faster. Indeed, we could have also saved ourselves a bit of time if we would have noticed that, since 13 is prime, the element 3 has a multiplicative inverse (in $\mathcal{U}(\mathbb{Z}_{13})$). A different way of saying that 3 has a multiplicative inverse is to say that 1 is a multiple of 3. Hence the order of 3 is greater than or equal to 1 (by Exercise 5.26). Conversely, 3 is obviously a multiple of 1 (it is $3 \cdot 1$) and so the order of the two elements match. Since the order 1 in $\mathbb{Z}_{13}$ is 13, this is also the order of 3. to prove the following theorem (as usual, p is a positive prime integer).

THEOREM 5.42. *Every non-zero element of $\mathbb{Z}_p$ has order p.*

PROOF. Let $[\,a\,]_p \in \mathbb{Z}_p$ be arbitrary. Like we just did above, since p is prime, $[\,a\,]_p$ has a has a multiplicative inverse $[\,b\,]_p$. Thus $b \cdot [\,a\,]_p = [\,1\,]_p$, i.e. 1 can be written as a multiple of $[\,a\,]_p$. On the other hand, $[\,a\,]_p = a \cdot [\,1\,]_p$ and hence it can be written as a multiple of 1. Thus by Exercise 5.26 (the one we just referenced in the example) the two elements have the same order. But since the order of $[\,1\,]_p$ is p (by a previous example), this is the order of $[\,a\,]_p$. Since $[\,a\,]_p \in \mathbb{Z}_p$ was arbitrary, every element of $\mathbb{Z}_p$ has order p. ■

EXERCISE 5.28. Prove that the corresponding theorem is always false when we consider a composite. In other words, prove that for all non-prime integers $n > 1$, there exists a non zero element $[\,a\,]_n \in \mathbb{Z}_n$ whose order is not n.

SOLUTION

REMARK 5.43. It's rather surprising that of all the groups we study, these $\mathcal{U}(\mathbb{Z}_n)$ groups, all be them abelian, give us the most trouble doing element order calculations. For just about every other class of examples we come across, $\mathcal{S}_n$, $\mathcal{D}_n$, $\mathbb{Z}_n$ (under addition of course), $\mathbf{SL}_2(\mathbb{Z})$, $\mathbb{S}^3$, and so on, we will be able to come with very simple formulas, often one can do in their heads, to find the order any element that can be confronted with. But even in simple cases, finding the order of an element of $\mathcal{U}(\mathbb{Z}_n)$ requires a boring multiplication. In hindsight, we shouldn't be too terribly surprised by this because we had a fairly difficult time finding the inverse of an element of this group as well!

EXAMPLE 5.44.

(1) Consider the matrix $\mathbf{A}$ which we set equal to

$$\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}.$$

Then we have that

$$\mathbf{A}^1 = \mathbf{A} \qquad \mathbf{A}^2 = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \qquad \mathbf{A}^3 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -\mathbf{I}_2$$

$$\mathbf{A}^4 = \mathbf{A}^3 \cdot \mathbf{A} = -\mathbf{A} \qquad \mathbf{A}^5 = -\mathbf{A}^2 \qquad \mathbf{A}^6 = \mathbf{I}_2.$$

Hence $|\,\mathbf{A}\,| = 6$. In fact, if $\mathbf{A}$ is any 2 by 2 matrix with determinant 1 and trace 1, using the Cayley–Hamilton Theorem from Exercise 3.91 on page 201, we have that

$$\mathbf{A}^2 = \mathbf{A} - \mathbf{I}_2.$$

Therefore

$$\mathbf{A}^3 = \mathbf{A} \cdot (\mathbf{A} - \mathbf{I}_2) = \mathbf{A}^2 - \mathbf{A} = (\mathbf{A} - \mathbf{I}_2) - \mathbf{A} = -\mathbf{I}_2.$$

As such, any such matrix satisfies $\mathbf{A}^6 = \mathbf{I}_2$. As long as we are not in $\mathbb{Z}_2$ this implies that such a matrix has order 6. (In $\mathbb{Z}_2$ since $-1 = 1$ such a matrix has order 3).

DEFINITION 5.45. ${}^{\pm}\mathbf{SL}_n(R)$ is defined to the set of 2 by 2 matrices with determinant $+1$ or -1.

EXERCISE 5.29.

SOLUTION

(1) Prove that for every integer n, ${}^{\pm}\mathbf{SL}_n(R)$ is a group.

Now suppose R be any commutative ring without zero divisors that we have been considering except for $\mathbb{Z}_2$. (These are $\mathbb{Z}_p$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ where p is an odd prime number.)

(2) Prove that if $\mathbf{A} \in {}^{\pm}\mathbf{SL}_2(R)$ has order 2 and determinant 1, then $\mathbf{A}$ is the matrix

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(3) Suppose that $\mathbf{A}$ has trace 0 and determinant $+1$, prove that $\mathbf{A}$ has order 4.

(4) Prove that if $\mathbf{A}$ has trace 0 and determinant -1, then $\mathbf{A}$ has order 2.

(5) Suppose that $\mathbf{A}$ has trace -1 and determinant $+1$, prove that $\mathbf{A}$ has order 3.

(6) Suppose that $\mathbf{A}$ has trace $+1$ and determinant $+1$, prove that $\mathbf{A}$ has order 6.

EXERCISE 5.30. Suppose that $\mathbf{A}$ has trace $+1$ and determinant -1.

SOLUTION

(1) Prove that $\mathbf{A}^n$ satisfies

$$\mathbf{A}^n = F_n \cdot \mathbf{A} + F_{n-1} \cdot \mathbf{I}_2$$

where F_n is the n^{th} term in the Fibonacci sequence first considered in Exercise 3.22 back on page 101.
Hint! Don't waste you time with this if you have done this already in a more general context somewhere else!

Now assume that $\mathbf{A}$ has at least one non-diagonal entry equal to ± 1. III

(2) Prove that the order $\mathbf{A}$ with $\mathbb{Z}_n$ coefficients is the first m so that both $F_m \equiv_n 0$ and $F_{m-1} \equiv_n +1$.

(3) Prove that the order $\mathbf{A}$ with $\mathbb{Z}_n$ coefficients is the first m so that both $F_m \equiv_n 0$ and $F_{m-2} \equiv_n -1$.

(4) Prove that the order $\mathbf{A}$ with $\mathbb{Z}_n$ coefficients is the first m so that both $F_{m-1} \equiv_n 1$ and $F_{m-2} \equiv_n -1$.

(5) Use the previous part and the table appearing at the end of the text on page 887 to calculate the order of $\mathbf{A}$ when coefficients are taken in

(a) $\mathbb{Z}_3$ **(b)** $\mathbb{Z}_5$ **(c)** $\mathbb{Z}_7$ **(d)** $\mathbb{Z}_{11}$

Careful! If the order you end up getting is odd you messed something up! (Why?)

Exercise 5.31. Let ${}^{\pm}\mathbf{SL}_2(\mathbb{Z})$ be as in Definition 5.45. Suppose that $\mathbf{A} \in {}^{\pm}\mathbf{SL}_2(\mathbb{Z})$ has finite order n. Prove that $n = 1, 2, 3, 4$, or 6. (So the only 2 by 2 matrix with integer coefficients whose 5^{th} power is equal to the identity is the identity)
Hint! Such a matrix, having integer entries, has integer trace. Cite an exercise from the chapter on matrices which restricts the trace of $\mathbf{A}$ to be between -2 and 2 with the cases that the trace is -2 or 2 occurring only in trivial circumstances. Show that if $\det \mathbf{A} = -1$ and $\operatorname{tr}(\mathbf{A}) \neq 0$, then $\operatorname{tr}\left(\mathbf{A}^2\right)$ is strictly larger than 2.

SOLUTION

Exercise 5.32. In what might seem like an exercise that doesn't fit in at all with the current flow, show that $|\mathcal{U}(\mathbb{Z}_n)| \leq 2$ iff $n = 1, 2, 3, 4$ or 6. hmmmm ... [IV]
Hint! Explicitly play with the Totient function defined in Remark 3.4.3 on page 153.

SOLUTION

Remark 5.46. It turns out that $|\mathcal{U}(\mathbb{Z}_n)|$ is the smallest degree an integer matrix of order n. For instance, since $\vartheta(5) = 4$, this means that there are also no integer 3 by 3 matrices of order 5, but there is a 4 by 4 one. Indeed, one can check directly that

$$\begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

is such an example (see Exercise 5.33)). More generally, since for any prime p, $\vartheta(p) = p - 1$, this fact guarantees us that for $m < p - 1$, there is not m by m integer matrix of order p but there is a $p - 1$ by $p - 1$ one (again, see Exercise 5.33 for the easy direction).

Exercise 5.33. Show that the $p - 1$ by $p - 1$ integer matrix

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -1 \\ 1 & 0 & 0 & \cdots & 0 & -1 \\ 0 & 1 & 0 & \cdots & 0 & -1 \\ 0 & 0 & 1 & \cdots & 0 & -1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{pmatrix}$$

has order p.

SOLUTION

[III] For instance, the *golden matrix*

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

always works.
[IV] See the remark, Remark 5.46, after the exercise.

Exercise 5.34. In contrast with Exercise 5.31, for any positive integer n show there exists a matrix $\mathbf{A} \in \mathbf{SL}_2(\mathbb{R})$ with $|\mathbf{A}| = n$.
Hint! Figure out how to represent a rotation by an angle $\frac{2\pi}{n}$ as a real 2 by 2 matrix.

Solution

Exercise 5.35.

Solution

(1) This exercise will require a set theoretical lemma. Let $f : X \to X$ be a function. A point $x \in X$ is a *fixed point* of f if $f(x) = x$. Suppose that

$$\sigma : X \to X$$

is an involution (that is it's a function with $\sigma \circ \sigma = \mathbf{1}_X$) and that $\#X$ is finite and odd. Prove that σ fixes at least one point.[V]
Hint! First convince yourself this is true. This should be the easy part. Then figure out the best way to right this up.[VI]

(2) Show that if G is a group of finite even order, then G contains an element of order 2.
Hint! Prove that if otherwise, the function

$$\begin{aligned} {}^{-1} &: G \setminus \{e_G\} \to G \setminus \{e_G\} \\ &: g \mapsto g^{-1} \end{aligned}$$

would be an involution between sets with an odd number of elements without a fixed point.

(3) Show that the converse is also true. That is prove that if G has odd order, then G does not contain an element of order 2.
Hint! Prove that if $s \in G$ has order 2, then

$$\begin{aligned} \mathrm{L}_s &: G \to G \\ &: g \mapsto sg \end{aligned}$$

is an involution without any fixed points.

Remark 5.47. Corollary 10.132 coming up on page 586 generalizes the above exercise, Exercise 5.35, to any prime p, not just the prime 2.

(2) Consider

$$\omega \stackrel{\text{def}}{=} \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}$$

discussed on page 128. There we showed that $\omega^6 = 1 = e_{\mathbb{H}^\times}$ and no lower power of ω is also equal to the identity. Therefore, by definition, ω has order 6.

[VI] Obviously this is false if we do not assume that $\#X$ is odd. For instance, if we take $X = \{\pm 1\}$ and define $\sigma(a) = -a$ (so $\sigma(1) = -1$ and $\sigma(-1) = 1$) then σ does not fix any points (since it doesn't fix 1 or -1).

(3) Consider

$$\mathfrak{L} := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}).$$

Then it is easy to check that

$$\mathfrak{L}^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$$

(indeed this is a an easy special case of a previous exercise since the trace of $\mathfrak{L}$ is 2 and the determinant is 1) which doesn't equal the identity matrix, $\mathbf{I}_2$, for any positive integer n. Thus $\mathfrak{L}$ has infinite (multiplicative) order.

(4) Instead, consider the $\mathbb{Z}_m$ version of $\mathfrak{L}$,

$$\mathfrak{L}_m := \begin{pmatrix} [1]_m & [1]_m \\ [0]_m & [1]_m \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}_m).$$

As in the previous example, we see that

$$\mathfrak{L}_m^n = \begin{pmatrix} [1]_m & [n]_m \\ [0]_m & [1]_m \end{pmatrix}.$$

Thus $\mathfrak{L}_m^n = \mathbf{I}_2$ iff $m \mid n$. The least positive integer where this is true is $m = n$. Hence the matrix $|\mathfrak{L}_m| = m$.

(5) The function

$$f(z) = \frac{z-1}{z} \in \mathbf{PGL}_2(\mathbb{R})$$

has order 3 (the operation is composition) since [VII]

$$(f \circ f)(z) = f(f(z)) = \frac{\frac{z-1}{z} - 1}{\frac{z-1}{z}} = \frac{\frac{z-1-z}{z}}{\frac{z-1}{z}} = \frac{1}{1-z}$$

$$(f \circ f \circ f)(z) = f(f(f(z))) = \frac{1}{1 - \frac{z-1}{z}} = \frac{z}{z-z+1} = z.$$

(6) The only element of $\mathbb{R}^+$ which have finite (multiplicative) order is 1. Indeed, if $r < 1$, then r^n determines a decreasing sequence

$$\cdots < r^3 < r^2 < r < 1$$

and if $r > 1$, then instead we have

$$\cdots > r^3 > r^2 > r > 1.$$

and neither could ever possibly equal the identity, which is here is 1. Similarly, the only element of $\mathbb{R}^\times$ which have finite order are $\{+1, -1\}$.

Recall from page 97 that

$$\{\{n\}\} \stackrel{\text{def}}{=} \left\{ e^{\frac{2\pi i m}{n}} : m \in \mathbb{Z}, 0 \le m < n \right\}.$$

EXERCISE 5.36. Suppose that $z \in \mathbb{C}^\times$ has finite order. Prove that $z \in \{\{n\}\}$ where n is the order of z. (In particular every element of finite order in $\mathbb{C}^\times$ is in $\mathbb{S}^1$.)
Hint! This is obvious if you know which theorem to use!

SOLUTION

[VII] For functions such as this we need to be careful when talking about powers since f^2 could easily mean $f \cdot f$ or $f \circ f$. Indeed, it is unfortunate that in a basic calculus course the two notations get fixed up. For instance, $\sin^{-1}(x)$ typically means the arc sine of x, not $\frac{1}{\sin x}$ while $\sin^2 x$ means $\sin x \cdot \sin x$ **not** $\sin(\sin(x))$. This is totally inconsistent and quite unfortunate!

(7) As an application of the previous example, we show that if $\mathsf{q} \in \mathbb{H}^\times$ has finite order, then $\mathsf{q} \in \mathbb{S}^3$ (i.e. has length 1). Now, since we have that

$$\| \mathsf{q}^n \| = \| \mathsf{q} \|^n$$

for any quaternion, if q has finite order n, then $1 = \| \mathsf{q}^n \| = \| \mathsf{q} \|^n$ which implies that $\| \mathsf{q} \| \in \mathbb{R}^+$ must also have finite order. But we just showed that the only such element is the number 1. Therefore we have that $\| \mathsf{q} \| = 1$, i.e. $\mathsf{q} \in \mathbb{S}^3$.

(8) The converse of the last example is false. In other words there are unit quaternions (i.e. elements of $\mathbb{S}^3$) which have infinite order. For example, see the below exercise.

EXERCISE 5.37. (Cayley Hamilton for 3 by 3 matrices) The *stagger* of a 3 by 3 matrix[VIII]

SOLUTION

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

is defined to be

$$\mathrm{sg}(\mathbf{A}) = (A_{11} A_{22} - A_{12} A_{21}) + (A_{11} A_{33} - A_{13} A_{31}) + (A_{22} A_{33} - A_{23} A_{32}).$$

Prove by hand, print out something from `MAPLE`, or site a reference which shows that

$$\mathbf{A}^3 = \mathrm{tr}(\mathbf{A})\, \mathbf{A}^2 - \mathrm{sg}(\mathbf{A})\, \mathbf{A} + \det(\mathbf{A})\, \mathbf{I}_3.$$[IX]

EXERCISE 5.38. Let $\mathbf{A}$ be the matrix

SOLUTION

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Calculate the order of $\mathbf{A}$ where the coefficients are taken in

(1) $\mathbb{Z}_2$, **(2)** $\mathbb{Z}_3$, **(3)** $\mathbb{Z}$.

(Maybe you want to use a computer to help do the multiplication more quickly.)

[VIII] It is the trace of the adjoint of $\mathbf{A}$. So when $\mathbf{A}$ is invertible, could also be seen as $\det \mathbf{A} \cdot \mathrm{tr}\mathbf{A}^{-1}$. Visually it is the determinant of the upper left 2 by 2 matrix plus the determinant of the lower right two by two matrix plus the determinant of the two by two matrix of the four corners.

[IX] This fits into a general statement which says that any n by n matrix is a root of its characteristic polynomial which is defined as

$$\det(\mathbf{A} - t\mathbf{I}).$$

The proof involves more about the theory of matrices that we need to go into. The case that $n = 2$–which was proven earlier–and the case $n = 3$–which is point of this problem–are the only two special cases we will be dealing with.

Exercise 5.39. Recall that the golden ratio is defined to be the number

$$\phi \stackrel{\text{def}}{=} \frac{1}{2} + \frac{\sqrt{5}}{2}.$$

More importantly, it satisfies the property that it's square is equal to 1 plus itself, that is

$$\phi^2 = \phi + 1.$$

The calculations done in this exercise may make the next exercise a bit less time consuming.

(1) Prove that $\phi^{-1} = \phi - 1 = -\widetilde{\phi}$ where $\widetilde{\phi}$ is the conjugate of ϕ, i.e. the other root of the equation $x^2 = x + 1$.

(2) Prove that ϕ^{-1} satisfies $x^2 + x = 1$ (so that squaring it adds 1 to its negative).

(3) Prove that $\phi^{-n} = (-1)^n \cdot (F_{n+1} - F_n\phi)$ where F_n is the n^{th} term in the Fibonacci sequence.

SOLUTION

Exercise 5.40. Calculate the order of the matrix

$$\frac{1}{2}\begin{pmatrix} 1 & \phi^{-1} & -\phi \\ \phi^{-1} & \phi & 1 \\ \phi & -1 & \phi^{-1} \end{pmatrix}.$$

Hint! If you like Cayley-Hamilton as much as I do, show that the trace and stagger are equal to ϕ while the determinant is 1 then use the 3 by 3 case of Cayley–Hamilton to rapidly take powers.

SOLUTION

Remark 5.48. An interesting way of doing Exercise 5.38 is to note that for the matrix $\mathbf{A}$ from that problem that

$$\mathbf{A}^3 = \mathbf{A} + \mathbf{I}_3$$

(where $\mathbf{I}_3$ is the 3 by 3 identity matrix). Therefore we have that

$$\begin{aligned}
\mathbf{A}^4 &= \mathbf{A} \cdot (\mathbf{A} + \mathbf{I}_3) = \mathbf{A}^2 + \mathbf{A}, \\
\mathbf{A}^5 &= \mathbf{A}^3 + \mathbf{A}^2 = \mathbf{A} + \mathbf{I}_3 + \mathbf{A}^2 \\
\mathbf{A}^6 &= \mathbf{A}^2 + \mathbf{A} + \mathbf{A}^3 = \mathbf{A}^2 + \mathbf{A} + \mathbf{A} + \mathbf{I}_3 = \mathbf{A}^2 + 2\mathbf{A} + \mathbf{I}_3 \\
\mathbf{A}^7 &= \mathbf{A}^3 + 2\mathbf{A}^2 + \mathbf{A} = \mathbf{A} + \mathbf{I}_3 + 2\mathbf{A}^2 + \mathbf{A} = 2\mathbf{A}^2 + 2\mathbf{A} + \mathbf{I}_3.
\end{aligned}$$

Taking everything mod 2 shows that $\mathbf{A}^7$ is the identity. Since 7 is prime, this implies that $\mathbf{A}$ has order 7. To figure out $\mathbf{A}^{13}$ note that

$$\begin{aligned}\mathbf{A}^{13} = \mathbf{A}^7 \cdot \mathbf{A}^6 &= (2\,\mathbf{A}^2 + 2\,\mathbf{A} + \mathbf{I}_3) \cdot (\mathbf{A}^2 + 2\,\mathbf{A} + \mathbf{I}_3) \\ &= (2\,\mathbf{A}^4 + 4\mathbf{A}^3 + 2\mathbf{A}^2) + (2\mathbf{A}^3 + 4\mathbf{A}^2 + 2\mathbf{A}) + (\mathbf{A}^2 + 2\mathbf{A} + \mathbf{I}_3) \\ &= 2\mathbf{A}^4 + 6\mathbf{A}^3 + 7\mathbf{A}^2 + 4\mathbf{A} + \mathbf{I}_3 \\ &= 2(\mathbf{A}^2 + \mathbf{A}) + 6(\mathbf{A} + \mathbf{I}_3) + 7\mathbf{A}^2 + 4\mathbf{A} + \mathbf{I}_3 \\ &= 9\mathbf{A}^2 + 12\mathbf{A} + 7\mathbf{I}_3.\end{aligned}$$

Therefore, $\mathbf{A}^{13}$ is equal to the identity when we take coefficients in $\mathbb{Z}_3$.

EXERCISE 5.41. Show that every non-identity matrix of the form

$$\mathbf{A} = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \tag{4}$$

with coefficients in $\mathbb{Z}_3$ has order 3 (by a previous theorem it's enough to show that $\mathbf{A}^3 = \mathbf{I}_3$, which theorem?).[x]

SOLUTION

This exercise shows that the previous exercise generalizes to all odd primes (and not just 3). It's proof might end up being different from the proof from the previous exercise. Of course if you can do this exercise, you do not need to do the previous one!

EXERCISE 5.42. Let $\mathbf{A}$ be a matrix in the form (4), but taken to have coefficients in $\mathbb{Z}_p$ – where p is an odd prime. Prove that $\mathbf{A}$ has order p.
Hint! Do an induction to show that if the coefficients are taken to be integers that

$$\mathbf{A}^n = \begin{pmatrix} 1 & na & nb + \frac{n(n-1)}{2}ac \\ 0 & 1 & nc \\ 0 & 0 & 1 \end{pmatrix}.$$

Then explain why each non-diagonal term must be divisible by n whenever $n = p$ is an odd prime. (Again, you only need to show that $\mathbf{A}^p = \mathbf{I}_3$! This **only!!** works since p is a prime!)

SOLUTION

This next exercise shows us that we can tell the order of most quaternions by inspection.

EXERCISE 5.43.

SOLUTION

[x] The set of such matrices with coefficients in a ring R is denoted $UT_3(R)$. It is the set of (3 by 3) upper triangle matrices with 1's down the diagonal. It is called the *Heisenberg group* .

(1) Let $\mathsf{q} \in \mathbb{S}^3$ and write it according to Theorem 3.103 as $\mathsf{q} = \cos(\theta) + \sin(\theta)\,\widehat{\mathbf{u}} \in \mathbb{S}^3$ ($\widehat{\mathbf{u}} \in \mathbb{S}^2$). Prove that q has finite order iff θ is a rational multiple of π.
Hint! use the fact that

$$\mathsf{q}^n = \cos(n\theta) + \sin(n\theta)\,\widehat{\mathbf{u}}.$$

So if $\mathsf{q}^n = 1$, then $n\theta = 2\pi m$ for some integer m.

(2) Suppose $\mathsf{q} \in \mathbb{S}^3$ and that

$$\mathbf{Re}\,\mathsf{q} = \cos\frac{2m\pi}{n}$$

with $\frac{m}{n}$ a reduced fraction (i.e. m and n have no common factors to cancel, i.e. they are relatively prime). Prove that q has order n.

EXAMPLE 5.49. For an example of a use of the previous exercise, Exercise 5.43, consider

$$\mathsf{q} = \frac{\sqrt{3}}{2} + \frac{1}{6}\mathbf{i} - \frac{1}{3}\mathbf{j} + \frac{1}{3}\mathbf{k}.$$

Then

$$\|\mathsf{q}\|^2 = \frac{3}{4} + \frac{1}{36} + \frac{1}{9} + \frac{1}{9} = \frac{27}{36} + \frac{1}{36} + \frac{4}{36} + \frac{4}{36} = 1$$

and so $\mathsf{q} \in \mathbb{S}^3$. Moreover we also have that

$$\mathbf{Re}\,\mathsf{q} = \frac{\sqrt{3}}{2} = \cos\frac{\pi}{6} = \cos\frac{2\pi}{12}.$$

Therefore $|\mathsf{q}| = 12$.

EXERCISE 5.44. Show that the element

$$\frac{4}{5} + \frac{3}{5}\mathbf{i}$$

is in $\mathbb{C}^\times$, $\mathbb{S}^1$, $\mathbb{H}^\times$, and $\mathbb{S}^3$. Show that it has infinite order.
Hint! You can use the fact that $\cos^{-1}(4/5)$ is an irrational multiple of π.
Bonus! Either prove that that $\cos^{-1}(4/5)$ is an irrational multiple of π or come up with a direct proof that the given element has infinite order. (Coming up with this direct proof isn't hard, it just takes a bit of focus. Prove that the powers of 5 in the denominators of powers of the given element keep going up by showing this to be the case for powers of two of this element!)

SOLUTION

EXERCISE 5.45. What's the order of the quaternion q where q is

SOLUTION

(1) $\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j}$

(2) $\frac{1}{2} + \frac{5}{6}\mathbf{i} + \frac{1}{6}\mathbf{j} - \frac{1}{6}\mathbf{k}$

(3) $\frac{1}{2}(\phi + \phi^{-1}\mathbf{i} + \mathbf{j})$

(4) $\frac{1}{2}(\phi\mathbf{i} + \phi^{-1}\mathbf{j} + \mathbf{k})$

(5) $\frac{\sqrt{2}}{4}\left((1-\sqrt{3}) + \frac{1}{3}\left[(1+\sqrt{3})\,\mathbf{i} + (4+2\sqrt{3})\,\mathbf{j} + 2\,\mathbf{k}\right]\right)$

Hint! After having fun showing $\|\mathfrak{q}\| = 1$, perhaps (ok, almost no sane person) you still don't recognize $\mathbf{Re}\,\mathfrak{q} = \frac{\sqrt{2}}{4}(1-\sqrt{3})$ as the cos of an angle. Consider instead $\mathbf{Re}\,\mathfrak{q}^2 = 2\,\mathbf{Re}\,\mathfrak{q}^2 - 1$ to make life easier. Alternatively, you may realize

$$\frac{\sqrt{2}}{4}(1-\sqrt{3}) = \frac{\sqrt{2}}{2}\cdot\frac{1}{2} - \frac{\sqrt{2}}{2}\cdot\frac{\sqrt{3}}{2} = \cos\theta\cdot\cos\psi - \sin\theta\cdot\sin\psi = \cos(\theta+\psi).$$

for some angles θ and ψ.

(6) $\frac{1}{2}(\phi + \phi^{-1}\mathbf{i})$

(7) $\frac{\sqrt{2}}{2}(1+\mathbf{i})$.

(8) $\mathbf{i} + \mathbf{j}$

where, as usual, ϕ is the golden ratio. (Don't forget to first check whether or not these dudes are actually elements of $\mathbb{S}^3$!)

EXERCISE 5.46. For a quaternion $\mathfrak{q}$ prove that

(1) $|\mathfrak{q}| = 2$ iff $\mathfrak{q} = -1$,

(2) $|\mathfrak{q}| = 4$ iff $\mathfrak{q} \in \mathbb{S}^2$.

SOLUTION

REMARK 5.50. If you prefer random analysis facts (to the random algebra fact I just spit out), we can see that there are an uncountable number of elements of $\mathbb{S}^3$ which

have infinite order. Indeed, an element

$$\cos\theta + \sin\theta\,\widehat{\mathbf{u}}$$

has finite order only when θ is a rational multiple (ranging from 0 to 2π) of π. However the irrational multiples of π are dense in that interval (in any interval). For instance, $\cos(\pi^2) + \sin(\pi^2)\mathbf{i}$ has infinite order since π is irrational and hence π^2 is an irrational multiple of π. (Of course we haven't and won't prove that π is an irrational number. This is actually quite difficult to prove.)

While elements of finite order may seem rare, they are actually quite common in group theory. In particular, we are about to see that every element of a finite group has finite order. This is particularly nice due to the fact that we will be studying, for the most part, finite groups.

PROPOSITION 5.51. *Let G be a group and suppose that $H \subseteq G$ satisfies*

(1) *H is closed under the operation from G;*

(2) *H is finite.*

Then every element of H has finite order.

PROOF. The result is vacuously true if H is empty. Therefore we assume that H is non-empty. Let $h \in H$ be arbitrary. Since H is closed under the operation, $h^n \in H$ for every positive integer n. If the set of all such powers

$$h, h^2, h^3, \ldots, h^n, \ldots$$

did not repeat, then this set would be infinite. But since it's contained inside of the finite set H it must repeat. That is there are distinct positive integers i and j so that $h^i = h^j$. For argument sake, we suppose that $i > j$. Multiplying both sides of the equality $h^i = h^j$ by h^{-j} we derived the equality

$$h^{i-j} = h^i h^{-j} = h^i (h^j)^{-1} = e_G.$$

Since $i - j > 0$ we have that h has finite order. ■

The intuition behind this proof says that if we keep taking powers of h, we gotta repeat.

COROLLARY 5.52. *Every element of a finite group has finite order.*

PROOF. A finite group satisfies the two properties of the above, Proposition 5.51. ■

THEOREM 5.53. *Suppose that $H \subseteq G$ is non-empty and satisfies the two properties of the Proposition 5.51. Then H is a group.*

PROOF. We verify the four group axioms.

Closure Closure holds by assumption.

Assoc. Associativity is inherited from the associativity of the operation in G.

Identity Since H is assumed to be non-empty, there exists some $h \in H$. By the previous proposition, Proposition 5.51, we know that h has finite order. That is $h^n = e_G$ for some positive integer n. But since H is closed every power of its element h is also in H. Therefore $e_G \in H$. Clearly e_G functions as an identity for the algebraic system H. This is because every element of H is an element of G, thus for any $h \in H$, $he_G = e_G h = h$.

Inverses Let $h \in H$ be arbitrary. We show that the inverse of h (which all we know by definition is that it lives in G) is a power of h and hence is an element of H by closure. Again, by the proposition h has finite order. Thus there exists some positive integer n so that $h^n = e_G$. But then

$$\begin{aligned} e_G &= h^n \\ &= h^{1+(n-1)} \\ &= h^1 h^{n-1} \\ &= h h^{n-1}. \end{aligned}$$

Similarly, we also have that $h^{n-1} h$ is also equal to the identity. Thus $h^{n-1} = h^{-1}$ is the inverse of h in G. Since H is closed under multiplication $h^{n-1} \in H$. Therefore $h^{-1} \in H$.[XI] ■

WARNING! 5.54. This is not true if H were not finite! For instance $\mathbb{N} \subseteq \mathbb{Z}$ is closed under the operation (here, that operation is addition) but $\mathbb{N}$ is not a group. On a similar note, $\mathbb{Z}^\times \subseteq \mathbb{Q}^\times$ is also closed under the operation (in this case the operation is the usual multiplication of rational numbers), but again $\mathbb{Z}^\times$ is not a group.

EXERCISE 5.47. (A bit harder.) Let p be a positive prime number different than 2. Show that 2 is a perfect square in $\mathbb{Z}_p$[XII] iff there exists a matrix in $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ whose order is 8.
Hint! We know what we need for a matrix to have order 4 (trace 0) and if a matrix has order 8 then its square has order 4.
SOLUTION

EXERCISE 5.48. For each integer n, compute $\mathbf{o}_n(\mathbb{Z}_4 \times \mathbb{Z}_4)$.
Hint! To make this go a bit faster note that every element of this group has order 1, 2, or 4. (So for numbers other than these your answer is 0.)There is only one element of order 1, namely the identity, $(0, 0)$. Now figure out how many elements of order 2 there are and lastly figure out how many elements of order 4 by taking what's left over.
SOLUTION

EXERCISE 5.49. Calculate the number of elements of each order in $\mathcal{S}_3 \times \mathbb{Z}_3$.
SOLUTION

5.9.2 CYCLIC GROUPS

[XI] In more detail, we proved that H has an identity, e_G, and for every $h \in H$ there exists an $\widehat{h}$ so that

$$h\widehat{h} = \widehat{h}h = e_G$$

[XII] This is not the case in $\mathbb{Z}_5$ but it is the case in $\mathbb{Z}_7$ since there $4^2 = 16 = 2$. It turns out that this is true iff at least one of $p + 1$ or $p - 1$ are divisible by 8.

DEFINITION 5.55. Let g be an element of some group.

(1) The *cyclic group* generated by g, denoted $\langle g \rangle$, is the set of all powers of g. In symbols:

$$\langle g \rangle \stackrel{\text{def}}{=} \{g^m : m \in \mathbb{Z}\} \text{ or in an additive group } \{mg : m \in \mathbb{Z}\}.$$

(2) A group G is *cyclic* if it is equal to the cyclic group generated by some element in it. In this case we call that element a *generator* of G.

Symbolically,

$$G \text{ is cyclic } \Leftrightarrow \exists g \in G, G = \langle g \rangle.$$

EXAMPLE 5.56.

(1) For any integer n the group $\mathbb{Z}_n$ is cyclic with generator $[\,1\,]_n$.

(2) For any integer n the group $\{\{n\}\}$ is cyclic with generator χ_n.

(3) For any integer n, the group

$$\{mn : m \in \mathbb{Z}\} = n\mathbb{Z}$$

is cyclic with generator the integer n.

REMARK 5.57. We will see later, in Theorem 6.13 coming up on page 316, that $\langle g \rangle$ is itself always a group (so the name "cyclic group" doesn't suck too badly!). This is easy to see now, but we do it later so its proof is in the proper context. For now we simply assume the result.

In symbols, a group G is cyclic iff

$$\exists g \in G \text{ s.t. } \forall h \in G \; \exists n \in \mathbb{Z} \text{ s.t. } h = g^n.$$

Or if you're like me and prefer words, if every element of G is a power of some element in it. Cyclic groups are the easiest class of groups to study. For instance, $\mathbb{Z}$ is cyclic (generated by 1 or -1), more generally $m\mathbb{Z}$ is cyclic (generated by m or $-m$) and the groups $\mathbb{Z}_n$ are cyclic (with generator $[\,1\,]_n$). Moreover, since

$$g^a g^b = g^{a+b} = g^{b+a} = g^b g^a$$

cyclic groups are always abelian. However, there are plenty of abelian groups, for example $\mathbb{Z}_2 \times \mathbb{Z}_2$, which are not cyclic.

Here's a proposition that will be useful to prove the corollary which follows. (You don't really need it! To tell you the truth, I just wanted an excuse to present the proof here so we will need the more general proposition later to prove Theorem 7.34!)

PROPOSITION 5.58. *Suppose that $h, k \in G$*

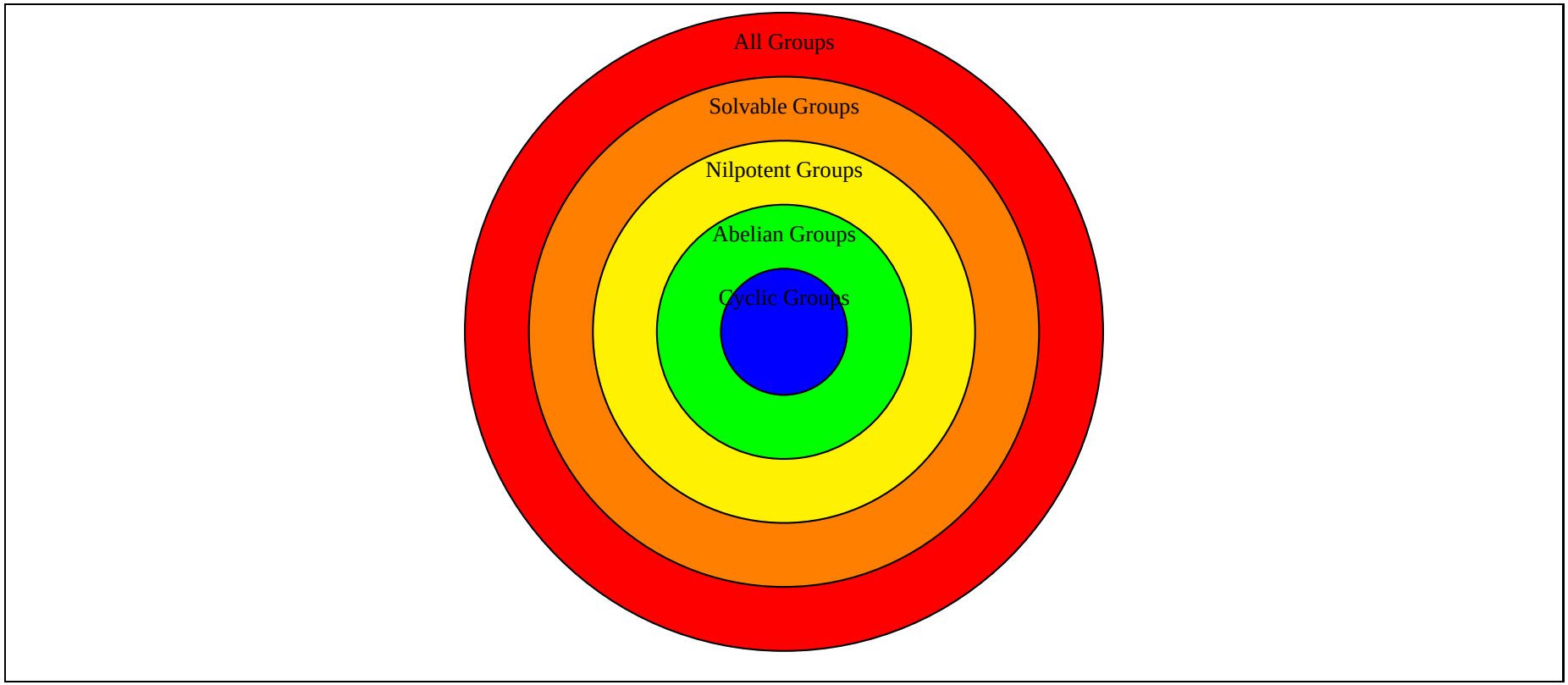


Figure 5.5: Types of Groups. We will define Solvable groups much later. Nilpotent groups will not be defined in these notes. They lie halfway between abelian groups and solvable groups. Cyclic groups are the most basic type of group.

(1) *Commute, and*

(2) *Satisfy* $\langle h \rangle \cap \langle \mathbf{k} \rangle = \{e_G\}$ [XIII]

Then $|h \cdot k| = \text{lcm}(|h|, |k|)$.

Proof. Let's denote $\text{lcm}(|h|, |k|)$ as $\mathcal{L}$. We need to show two things about $\mathcal{L}$:

- First we need to show that $(h \cdot k)^{\mathcal{L}} = e_G$. Since h and k commute we have that $(h \cdot k)^n = h^n \cdot k^n$ for all integers n. [XIV]. But since $|h| \Big| \mathcal{L}$, we have that $h^{\mathcal{L}} = e_G$. Also, since $|k| \Big| n$ we also have that $k^{\mathcal{L}} = e_G$. Therefore
$$(h \cdot k)^{\mathcal{L}} = h^{\mathcal{L}} \cdot k^{\mathcal{L}} = e_G \cdot e_G = e_G.$$
- Suppose that $(h \cdot k)^m = e_G$ where $m \in \mathbb{Z}^+$. We show that $\mathcal{L} \leq m$ by showing that $\mathcal{L} \mid m$. Since
$$(h \cdot k)^m = h^m \cdot k^m = e_G \Rightarrow h^m = (k^m)^{-1} = k^{-m}$$

[XIII] That is h and k have only the identity as a common power. That is the only solution to the equation $h^n = k^m$ is $h = k = e_G$.
[XIV] Don't forget, this never happens when a and b do not commute!!

we have that $h^m = k^{-m}$. Therefore as a power of h (it's written that way) and a power of k (since we just showed it to equal one) we have $h^m \in \langle h \rangle \cap \langle \mathbf{k} \rangle$. But the assumption, Assumption 2, says that only the identity can be in this intersection and therefore we have that $h^m = e_G$. common is assumed to be e_G, we must have $a^m = e_G$.

Moreover as k^{-m} is equal to h^m which we have just shown to be the identity, we also have that $k^{-m} = e_G$. Taking inverses of both sides of this last identity shows that, in addition to $h^m = e_G$ proven in the previous paragraph, we also have $k^m = e_G$ as well.

Therefore, by Theorem 5.38**(2)**, $|h| \Big| m$ and $|k| \Big| m$. Lastly, as $\mathcal{L}$ is defined to be the least common multiple of $|h|$ and $|k|$ (least common multiple divides any common multiple), we also have that $\mathcal{L} \mid m$ as desired. ■

COROLLARY 5.59. *For $(h,k) \in H \times K$, $|(h,k)| = \operatorname{lcm}(|h|, |k|)$.*

PROOF. It's trivial to see that the elements $\widetilde{h} := (h, e_K)$ and $\widetilde{k} := (e_H, k)$

(1) Commute;

(2) Satisfy

$$\left\langle \widetilde{h} \right\rangle \cap \left\langle \widetilde{k} \right\rangle = e_{H\times K}.$$

Therefore their product, which is

$$\widetilde{h}\,\widetilde{k} = (h, e_K)(e_H, k) = (h, k),$$

has order

$$\operatorname{lcm}(|\widetilde{h}|, |\widetilde{k}|) = \operatorname{lcm}(h, k)$$

by the proposition. ■

For instance, the element $(\rho, [\,5\,]_6) \in \mathcal{D}_4 \times \mathbb{Z}_6$ has order $\operatorname{lcm}(4,6) = 12$.

So far we have defined the order of a group (a set!) simply as the number of elements in it, and the order of an element of a group as the least positive power of that element which equals the identity. The following corollary connects these two different meanings of the word "order" we have come across so far. It says the order of an element of a group is the same as the order of the group generated by that element.

NOTATION 5.60. Let $m \in \mathbb{N}$ and g an element of some group. Define

$$\begin{aligned} \mathbf{E}_g^m &: \mathbb{Z}_m \to \langle g \rangle \\ &: [\,a\,]_d \mapsto g^a. \end{aligned}$$

In words, the following theorem states that the number of elements in a cyclic group is equal to the order of any generator of that cyclic group.

THEOREM 5.61. *The set $\langle g \rangle$ has $|g|$ elements in it.*

PROOF. Let's denote $|g|$ by n. Recall from Theorem 5.38**(3)** that $g^a = g^b$ holds iff $a \equiv_n b$. The current theorem is almost a direct translation of that fact. In more detail, we show that the map

$$\begin{aligned} \mathbf{E}_g^n &: \mathbb{Z}_n \to G \\ &: [\,a\,]_n \mapsto g^a \end{aligned}$$

just defined is a bijection. We show that $\mathbf{E}_g^n$ is a bijection $\mathbb{Z}_n \to \langle g \rangle$. Since the former set has n elements in it, this will prove that so does the latter.

(1) (**Well defined**) Suppose that

$$a \equiv_n b.$$

Then we know by 5.38**(3)** that $g^a = g^b$. Hence $\mathbf{E}_g^n$ is well defined.

(2) (**One-to-one**) Suppose that $\mathbf{E}_g^n([\,a\,]_n) = \mathbf{E}_g^n([\,b\,]_n)$. Then by definition of $\mathbf{E}_g^n$ we have that $g^a = g^b$. But again, by 5.38**(3)** (using its oppositive implication) we now see that $a \equiv_n b$. Therefore $\mathbf{E}_g^n$ is an injection.

(3) (**Onto** $\langle\, g \,\rangle$) We need to prove the equality of two sets:

$$\mathrm{Im}\left(\mathbf{E}_g^n\right) = \langle\, g \,\rangle.$$

$\subseteq$ By definition, for every $[\,a\,]_n \in \mathbb{Z}_n$, $\mathbf{E}_g^n([\,a\,]_n) = g^a$, a power of g. Therefore we have that $\mathrm{Im}\left(\mathbf{E}_g^n\right) \subseteq \langle\, g \,\rangle$.

$\supseteq$ Let $h \in \langle\, g \,\rangle$. Then, by definition of $\langle\, g \,\rangle$, this implies that there exists an integer m so that $h = g^m$. But then

$$\mathbf{E}_g^n([\,m\,]_n) = g^m = h.$$

Therefore we also have that $\mathrm{Im}\left(\mathbf{E}_g^n\right) \supseteq \langle\, g \,\rangle$.

This completes the proof that $\mathbf{E}_g^n$ is a bijection. ∎

We record the specific information about the map $\mathbf{E}_g^n$ defined in the proof.

COROLLARY 5.62. *If $n = |\,g\,|$, then $\mathbf{E}_g^n$ is a bijection.*

Actually, since if $n \mid m$ the map

$$\begin{aligned} \pi_n^m &: \mathbb{Z}_m \to \mathbb{Z}_n \\ &: [\,a\,]_m \mapsto [\,a\,]_n \end{aligned}$$

is well defined, we know that the composition, $\mathbf{E}_n^g \circ \pi_n^m$ is well defined. But this composition takes

$$[\,a\,]_m \mapsto g^a.$$

This is just the map

$$\mathbf{E}_g^m : \mathbb{Z}_m \to \langle\, g \,\rangle.$$

With a little more focus we can fill in the details of the following:

COROLLARY 5.63. *The relation*

$$\begin{aligned} \Pi_g^h &: \langle\, h \,\rangle \rightsquigarrow \langle\, g \,\rangle \\ &: h^a \rightsquigarrow g^a \end{aligned}$$

is well defined iff $|\,g\,| \Big|\, |\,h\,|$. Moreover it is always surjective. ∎

As a particular case, since $\Pi_g^{[\,1\,]_m} = \mathbf{E}_g^m$, $\mathbf{E}_g^m$ is well defined iff $|\,g\,| \Big|\, \overset{|[\,1\,]_m|}{\overset{\shortparallel}{m}}$.

Remark 5.64. Later, we will define what it means for two groups to be isomorphic. Roughly, it means that the multiplication tables line up. In that context, we will prove that the map $\mathbf{E}_g^n$ is an isomorphism of groups.

Exercise 5.50.

(1) Prove $\Pi_g^h \circ \Pi_h^k = \Pi_g^k$.

(2) Prove Corollary 5.63.

SOLUTION

COROLLARY 5.65. *A finite group G is cyclic iff there exists an element $g \in G$ whose order is $|G|$.*

PROOF.

$\Rightarrow$ Suppose that G is cyclic. Then by definition, G contains an element g with $G = \langle g \rangle$. But the order of the left hand side is $|G|$ and the order of the right hand side is the order of g by the theorem, Theorem 5.61.

$\Leftarrow$ Conversely, if g is an element of G whose order is the same as the order of G, we have that $\langle g \rangle$ and G have the same number of elements. But since we always have that $\langle g \rangle$ and G is assumed finite, we must have equality, $G = \langle g \rangle$. In conclusion, G is cyclic generated by g. ■

REMARK 5.66. Later, when we talk about isomorphisms, we will show that in fact these two sets are isomorphic, a condition much stronger than having the same number of elements.

EXAMPLE 5.67. If we consider the element $([1]_4, [1]_6) \in \mathbb{Z}_4 \times \mathbb{Z}_6$, which has order 12, the map

$$\mathbf{E}^{12}_{(1,1)} : \mathbb{Z}_{12} \to \mathbb{Z}_4 \times \mathbb{Z}_6$$

defined in the theorem, Theorem 5.61, behaves as follows:

$$\begin{array}{llll}
[1]_{12} \mapsto ([1]_4, [1]_7) & [2]_{12} \mapsto ([2]_4, [2]_7) & [3]_{12} \mapsto ([3]_4, [3]_7) & [4]_{12} \mapsto ([0]_4, [4]_7) \\
[5]_{12} \mapsto ([1]_4, [5]_7) & [6]_{12} \mapsto ([2]_4, [0]_7) & [7]_{12} \mapsto ([3]_4, [1]_7) & [8]_{12} \mapsto ([0]_4, [2]_7) \\
[9]_{12} \mapsto ([1]_4, [3]_7) & [10]_{12} \mapsto ([2]_4, [4]_7) & [11]_{12} \mapsto ([3]_4, [5]_7) & [0]_{12} \mapsto ([0]_4, [0]_7).
\end{array}$$

EXAMPLE 5.68. Since $([1]_2, [1]_3)$ has order 6, $\mathbb{Z}_2 \times \mathbb{Z}_3$ is cyclic.

More generally, suppose that we have cyclic groups C_1 with $|C_1| = n$ and C_2 with order $|C_2| = m$, (for example, we can imagine $C_1 = \mathbb{Z}_n$ and $C_2 = \mathbb{Z}_m$).

If the orders m and n are relatively prime, then we have $\text{lcm}(n, m) = n \cdot m$, and so by Exercise 5.59, for any generators $x_1 \in C_1$ and $x_2 \in C_2$ we have (x_1, x_2) has order $n \cdot m = |C_1 \times C_2|$. Therefore, the group, having an element equal to its order, is cyclic.

Conversely, if n and m are **not** relatively prime, meaning $\gcd(n, m) > 1$, so the least common multiple of n and m, which is given by the fraction

$$\text{lcm}(n, m) = \frac{nm}{\gcd(n, m)} < nm.$$

Therefore (x_1, x_2) no longer has order $n \cdot m$. Could there be some other element that does have order $n \cdot m$? No, here's why. Let's denote this least common multiple by $\mathcal{L}$. Then we have for all $g_1 \in C_1$ and $g_2 \in C_2$,

$$\begin{aligned}
(g_1, g_2)^{\mathcal{L}} &= (g_1^{\mathcal{L}}, g_2^{\mathcal{L}}) \\
&= (e, e) \\
&= \text{The identity of } C_1 \times C_2.
\end{aligned}$$

Therefore, every element of $C_1 \times C_2$ has order less than or equal to $\mathcal{L}$, (actually, a divisor of $\mathcal{L}$) and hence strictly less than $n \cdot m$. As $C_1 \times C_2$ has order $n \cdot m$, this means that no element of the group $C_1 \times C_2$ has order equal to $C_1 \times C_2$ itself, and so, we have proven that in this case $C_1 \times C_2$ is not cyclic. Carrying out the exact same argument for a collection of cyclic groups, we have proven

THEOREM 5.69. *For finite cyclic groups $C_1, \ldots, C_r$ the group $C_1 \times \cdots \times C_r$ is cyclic iff the integers $|C_1|, \ldots, |C_r|$ are pair-wise relatively prime. Moreover in the cyclic cases, $(g_1, \ldots, g_r)$ is a generator for any generators $g_i \in C_i$.* ■

Here's an application of this theorem:

COROLLARY 5.70 (Chinese Remainder Theorem). *Suppose that $n_1, \ldots, n_r$ are pair-wise relatively prime integers. Then for any list of r integers, $a_1, \ldots, a_r$ there exists an integer x so that*

$$x \equiv_{n_1} a_1, \quad \ldots \quad , x \equiv_{n_r} a_r.$$

PROOF. For notational ease, we restrict ourselves to the case that $r = 2$ and denote a_1 as a and a_1 as b. The more general proof is the same after some dot, dot, dots have been added in.

Assume that n and m are relatively prime integers. Then, the ordered pair $([\,a\,]_n, [\,b\,]_m)$ is an element of $\mathbb{Z}_n \times \mathbb{Z}_m$. Since n and m are assumed to be relatively prime, the theorem says that this group, $\mathbb{Z}_n \times \mathbb{Z}_m$, is cyclic with generator $([\,1\,]_n, [\,1\,]_m)$. Therefore the element $([\,a\,]_n, [\,b\,]_m)$ is a power of this generator, $([\,1\,]_n, [\,1\,]_m)$. That is there exists an integer x so that

$$x \cdot ([\,1\,]_n, [\,1\,]_m) = ([\,a\,]_n, [\,b\,]_m).^{\text{XV}}$$

But

$$x \cdot ([\,1\,]_n, [\,1\,]_m) = \underbrace{([\,1\,]_n, [\,1\,]_m) + \cdots + ([\,1\,]_n, [\,1\,]_m)}_{x \text{ times}} = ([\,x\,]_n, [\,x\,]_m).$$

Therefore $[\,x\,]_n = [\,a\,]_n$ and $[\,x\,]_m = [\,b\,]_m$, i.e. $x \equiv_n a$ and $x \equiv_m b$, as desired. ■

EXAMPLE 5.71. (Leaving off brackets) Let's solve the simultaneous equivalences

$$x \equiv_4 2 \quad \text{and} \quad x \equiv_5 3.$$

Let's look at the powers of $(1, 1)$. The first multiple of $(1, 1)$ that equals $(2, 3)$ is the desired integer. Indeed, we have

$$\begin{aligned}
1 \cdot (1,1) &= (1,1) & 2 \cdot (1,1) &= (2,2) & 3 \cdot (1,1) &= (3,3) \\
4 \cdot (1,1) &= (0,4) & 5 \cdot (1,1) &= (1,0) & 6 \cdot (1,1) &= (2,1) \\
7 \cdot (1,1) &= (3,2) & 8 \cdot (1,1) &= (0,3).
\end{aligned}$$

Now we could keep playing this game, but we can end this early since we now know

$$\begin{aligned}
18 \cdot (1,1) &= (2 \cdot 5 + 8)(1,1) \\
&= 2 \cdot 5(1,1) + 8(1,1) \\
&= 2\,(1,0) + (0,3) \\
&= (2,0) + (0,3) \\
&= (2,3).
\end{aligned}$$

Indeed, as a check note that $18 \equiv_4 2$ and $18 \equiv_5 3$, as desired. Motivated by this, we improve the Chinese Remainder Theorem (CRT) a bit in the exercises below.

[XV] We're writing the abelian group $\mathbb{Z}_n \times \mathbb{Z}_m$ additively.

Exercise 5.51. Let $n_1, \ldots, n_r$ be a set of integers. Prove that

Solution

(1) this set of integers are pair-wise relatively prime iff

$$\operatorname{lcm}(n_1, \ldots, n_r) = n_1 \cdot \cdots \cdot n_r.$$[XVI]

(2) ⟨This part shows quite the opposite is true for gcd's ⟩ by way of example, it's possible for

$$\gcd(n_1, \ldots, n_r) = 1$$

without any pair n_i, n_j, $1 < i, j < r$ being relatively prime.
Hint! For this second part, why bother with anything besides $r = 3$?

In this next exercise, we present a bit more streamlined proof of the Chinese remainder theorem at least in the case we are dealing with just two integers.

Exercise 5.52. Assume that n and m are relatively prime integers. Let $\widehat{n}$ and $\widehat{m}$ be integers so that

Solution

$$n \cdot \widehat{n} \equiv_m 1 \quad \text{and} \quad m \cdot \widehat{m} \equiv_n 1.$$[XVII]

(1) Let $a, b \in \mathbb{Z}$ be arbitrary. Define $x = x(a, b) = a\widehat{n}\,n + b\widehat{m}\,m$. Prove that

$$x \equiv_n a \quad \text{and} \quad x \equiv_m b.$$

(2) Prove that

$$\widehat{n}\,n \cdot ([\,1\,]_n, [\,1\,]_m) = ([\,0\,]_n, [\,1\,]_m), \qquad \widehat{m}\,m \cdot ([\,1\,]_n, [\,1\,]_m) = ([\,1\,]_n, [\,0\,]_m)$$

and

$$(\widehat{n}\,n + \widehat{m}\,m)\,([\,1\,]_n, [\,1\,]_m) = ([\,1\,]_n, [\,1\,]_m).$$

These things exist by Theorem 3.139 back on page 151.

We will use this exercise in the example below.

Exercise 5.53. Let $[\,a\,]_n \in \mathcal{U}(\mathbb{Z}_n)$.

Solution

(1) Suppose that n happened to be prime and that $[\,a\,]_n$ had order 2. Prove that $[\,a\,]_n = [\,-1\,]_n$.

[XVI] In words the lcm of the set of integers is the product of that set.
[XVII] So $\widehat{n}$ is a multiplicative inverse of n mod m and $\widehat{m}$ is a multiplicative inverse of m mod n.

(2) Suppose that m is the first positive integer so that $[\,a\,]_n^m = -1 = [\,n-1\,]_n$. Prove that the order of $[\,a\,]_n$ is $2m$.

(3) Assume that $[\,a\,]_n$ has order greater than 2 and that $[\,a\,]_n^p = -1$ for some positive prime number p. Prove that the order of $[\,a\,]_n$ is $2p$.

Hint! What are the factors of $2p$? Why can't $[\,a\,]_n$ have order equal to any of these factors?

EXAMPLE 5.72. In this example we make implicit use of the above exercise, Exercise 5.53.

(1) The group $\mathcal{U}(\mathbb{Z}_{11})$ (which has order $11-1=10$) is cyclic since (leaving off brackets)

$$2^1 = 2 \qquad 2^2 = 2\cdot 2 = 4 \qquad 2^3 = 2\cdot 4 = 8$$
$$2^4 = 2\cdot 8 = 16 = 5 \qquad 2^5 = 2\cdot 5 = 10$$

which implies that $|\,2\,|\ (= \left|\,[\,2\,]_{10.}\,\right|) = 10$. Thus, there exists an element, namely 2, whose order matches the order of the group (which is 10), and the group is cyclic by Corollary 5.65.

(2) There are

$$\vartheta(15) = 15\cdot\left(1-\frac{1}{5}\right)\left(1-\frac{1}{3}\right) = 4\cdot 2 = 8.$$

elements in $\mathcal{U}(\mathbb{Z}_{15})$. They are

$$\{1,2,4,7,8,11,13,14\}.$$

I claim that $\mathcal{U}(\mathbb{Z}_{15})$ is not cyclic. Indeed

Element	1	2	4	7	8	11	13	14
Order	1	4	2	4	4	2	4	2

Therefore, no element of $\mathcal{U}(\mathbb{Z}_{15})$ has order $\vartheta(15)=8$. Therefore, $\mathcal{U}(\mathbb{Z}_{15})$ is not cyclic by Corollary 5.65. (After proving Corollary 5.88, we will be able to do this faster by simply noting that this group contains more than 1 element of order 2.)

(3) The group $\mathcal{U}(\mathbb{Z}_{18})$ is cyclic. For starters, since $18 = 3^2\cdot 2$, we have that the order of $\mathcal{U}(\mathbb{Z}_{18})$ is

$$\vartheta(18) = 18\cdot\left(1-\frac{1}{2}\right)\cdot\left(1-\frac{1}{3}\right) = 2\cdot 3 = 6.$$

Moreover, the order of the element $[\,5\,]_{18}$ has order 6 since

$$5^1 = 5 \qquad 5^2 = 25 = 7 \qquad 5^3 = 5\cdot 7 = 35 = -1$$
$$5^4 = 5\cdot -1 = -5 \qquad 5^5 = -25 = -7 \qquad 5^6 = 1.$$

REMARK 5.73. Again, we could have done this much faster. Simply by computing

$$5^3 = 125 = 7 \cdot 18 - 1 \equiv_{18} -1$$

we could have computed that $|[5]_{18}| = 6$.

(4) The group $\mathcal{U}(\mathbb{Z}_{27})$ is cyclic. We show this by finding an element of order

$$\vartheta(27) = 27 \cdot \left(1 - \frac{1}{3}\right) = 18.$$

I claim that $[2]_{27}$ does the job. Indeed, we just need to show

$$2^2 = 4 \neq 1 \qquad 2^3 = 8 \qquad 2^6 = 10 \qquad 2^9 = 10 \cdot 8 = 26 = -1.$$

This shows that the order of 2 is 18 (why?), as desired.

(5) So if $\mathcal{U}(\mathbb{Z}_n)$ is cyclic, is 2 a generator? The answer is no! For example, $\mathcal{U}(\mathbb{Z}_7)$ is cyclic, with generator 3 since $3^3 = 6 = -1$ which implies that $|3| = 6$ (by the Exercise), the order of $\mathcal{U}(\mathbb{Z}_7)$. However 2 is not a generator since $2^3 = 8 = 1 \bmod 7$, the order of 2 is 3, not 6.

(6) Ok, now we are getting desperate. Perhaps if $\mathcal{U}(\mathbb{Z}_n)$ is cyclic, then either 2 or 3 is a generator. This too is false and the first counter example comes for $n = 23$ (Jim Carey would be happy.) Indeed, there we have that the order of both 2 and 3 is 11, which doesn't match the order of the group, ($|\mathcal{U}(\mathbb{Z}_{23})| = 23 - 1 = 22$). Here, 5 (meaning $[5]_{23}$) is a generator though. Indeed,

$$\begin{aligned}
5^{11} &= 5^{2\cdot 5+1} \\
&= 5^{2\cdot 5} \cdot 5 \\
&= \left(5^2\right)^5 \cdot 5 \\
&= (25)^5 \cdot 5 \\
&\equiv_{23} (2)^5 \cdot 5 \\
&= 32 \cdot 5 \\
&\equiv_{23} 9 \cdot 5 \\
&= 45 \\
&\equiv_{23} -1
\end{aligned}$$

which implies that the (multiplicative) order of 5 (really I mean $[5]_{23}$) is 22, which is the order of $\mathcal{U}(\mathbb{Z}_{23})$. We could have done this more efficiently. Since $\mathcal{U}(\mathbb{Z}_{23})$ has order 22, the only possible orders of $\mathcal{U}(\mathbb{Z}_{23})$ are 1,2,11, and 22. As always for primes, the only element of order 2 is $[-1]_{23}$. Now taking powers of $[-1]_{22}$ we see

$$\begin{array}{llll}
2^1 = 2 & 2^2 = 4 & 2^3 = 8 & 2^4 = 16 \\
2^5 = 32 \equiv_{23} 9 & 2^6 \equiv_{23} 2 \cdot 9 = 18 & 2^7 \equiv_{23} 2 \cdot 18 = 36 \equiv_{23} 13 & 2^8 \equiv_{23} 2 \cdot 13 = 26 \equiv_{23} 3 \\
2^9 \equiv_{23} 2 \cdot 3 = 6 & 2^{10} \equiv_{23} 2 \cdot 6 = 12 & 2^{11} \equiv_{23} 2 \cdot 12 = 24 \equiv_{23} 1. &
\end{array}$$

Which implies, as we already stated, that the order $[\,2\,]_{23}$ in this group is 11. This fact alone could have been seen faster:

$$\begin{aligned} 2^{11} &= 2^6 \cdot 2^5 \\ &= 64 \cdot 32 \\ &\equiv_{23} (-5) \cdot 9 \\ &= -(45) \\ &\equiv_{23} -(-1) \\ &= 1, \end{aligned}$$

but we are more interested in the powers we got of 2. Since 11 is prime, all of the first 10 powers of 2 mod also have order 11. Moreover in a cyclic group of order 22, there are exactly 10 elements of order 22, so these are all of the elements of order 11. Thus the ten elements of order 11 in $\mathbb{Z}_{23}$ are

$$\{[\,2\,]_{23}, [\,4\,]_{23}, [\,8\,]_{23}, [\,16\,]_{23} = [\,-7\,]_{23}, [\,9\,]_{23}, [\,18\,]_{23} = [\,-5\,]_{23}, [\,13\,]_{23} = [\,-10\,]_{23}, [\,3\,]_{23}, [\,6\,]_{23}, [\,12\,]_{23} = [\,-11\,]_{23}\}.$$

From our general theory, there is 1 element of order 1, which must be $[\,1\,]_{23}$ and one element of order 2, which must be $[\,-1\,]_{23}$. Moreover, the remaining $\varphi(22) = 10$ elements have order 22 and are therefore generators. Incidentally, those remaining 10 elements are the negatives of the initial 10 elements and are

$$\{[\,21\,]_{23} = [\,-2\,]_{23}, [\,19\,]_{23} = [\,-4\,]_{23}, [\,15\,]_{23} = [\,-8\,]_{23}, [\,7\,]_{23}, [\,14\,]_{23} = [\,-9\,]_{23}, [\,5\,]_{23}, [\,20\,]_{23} = [\,-3\,]_{23}, [\,17\,]_{23} = [\,-6\,]_{23}, [\,11\,]_{23}\}.$$

REMARK 5.74. It turns out that if p is a prime number, then the group $\mathcal{U}(\mathbb{Z}_p)$ is cyclic (of order $p-1$). We give a "half proof" below in Theorem 5.75. However it's quite difficult in general (without a stroke of luck) to find a generator! Indeed, we've seen that for $p = 23$ neither 2 or 3 are generators, but 5 is. We can actually push this further, for $p = 41$ none of 2, 3, or 5 are generators, but 7 is. The first time none of 2, 3, 5, or 7 are generators occurs when $p = 109$. What I find equally strange is for the prime $p = 173$ all four, 2, 3, 5, and 7, are generators.

EXERCISE 5.54. Prove that the group $\mathcal{U}(\mathbb{Z}_{50})$ is cyclic.
Careful! Of course 50 isn't prime, so the theorem below, Theorem 5.75, is useless here.

SOLUTION

In Corollary E.28 on page 818 in the appendix we prove the following result:

THEOREM 5.75. *For any prime p, $\mathcal{U}(\mathbb{Z}_p)$ is cyclic of order $p-1$.* ■

Let's give you an exercise to show you that you are not getting the wrong idea about what I was doing in the above remark.

EXERCISE 5.55. Show that neither 2 nor 3 are generators of $\mathcal{U}(\mathbb{Z}_{103})$, but the element 6 is a generator.

SOLUTION

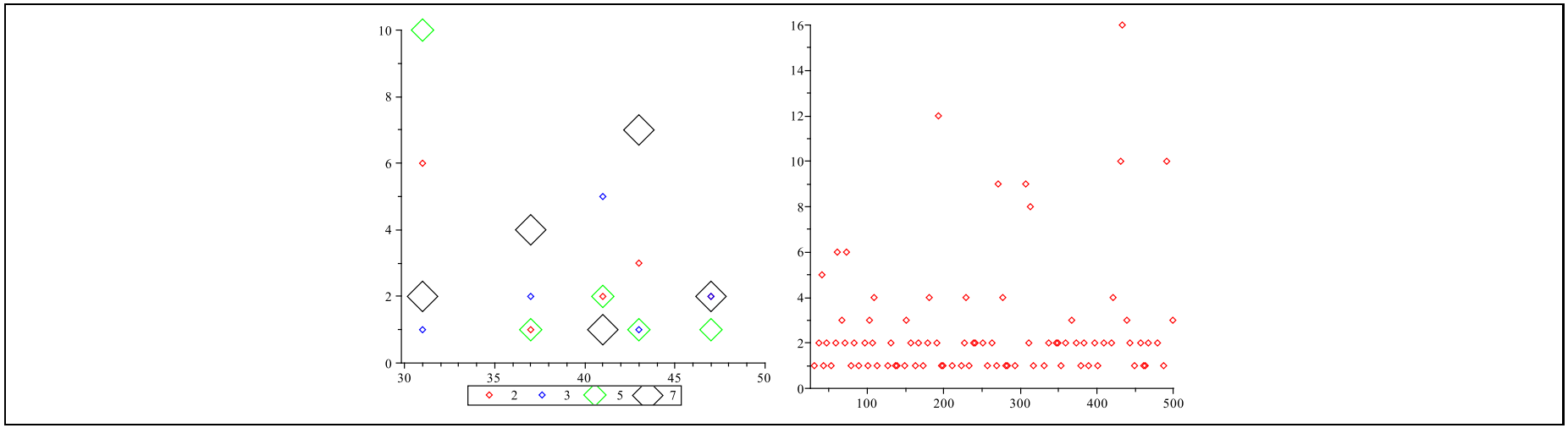


Figure 5.6: These MAPLE generated images show us how closely $[\,i\,]_p$ is to being a generator of $\mathcal{U}(\mathbb{Z}_p)$ for $i = 2$, 3, 5, and 7. More precisely, the image on the left plots p against $\frac{|[\,2\,]_p|}{p-1}$ in red, $\frac{|[\,3\,]_p|}{p-1}$ in blue, $\frac{|[\,5\,]_p|}{p-1}$ in green, and $\frac{|[\,7\,]_p|}{p-1}$ in **black** for all primes p between 30 and 50 (of course these ratios are 1 iff the element considered is a generator). For instance, the graph shows that of the four, only 7 is a generator of $\mathcal{U}(\mathbb{Z}_{41})$ and that this is the only p between 30 and 50 where 7 is a generator. It also shows that only 2 and 3 are generators of $\mathcal{U}(\mathbb{Z}_{37})$. The image on the right deals only with $i = 3$ and plots p vs $\frac{|[\,3\,]_p|}{p-1}$ for all primes p between 30 and 500.

EXERCISE 5.56. Find a generator of $\mathcal{U}(\mathbb{Z}_{109})$. (Feel free to use a calculator!) (note that 109 is a prime.)

SOLUTION

EXERCISE 5.57. Prove that the group of positive rational number under multiplication do not form a cyclic group.

SOLUTION

5.9.3 ORDERS IN CYCLIC GROUPS

Given the order of an element g, the next several results give us a formula to compute the order of any power of g. Throughout, we assume that g has finite order. We will be able to break the general case down into two subcases, the case that m divides the order of g and the case that m is relatively prime to the order of g.

LEMMA 5.76. *Suppose that g_1 and g_2 are group elements. Then*

$$\langle\, g_1 \,\rangle \subseteq \langle\, g_2 \,\rangle \Leftrightarrow g_1 \in \langle\, g_2 \,\rangle\,.$$

PROOF.

$\Rightarrow$ Suppose $\langle\, g_1 \,\rangle \subseteq \langle\, g_2 \,\rangle$. Then since $g_1 = g_1^1$ we know that $g_1 \in \langle\, g_1 \,\rangle$ which implies that $g_1 \in \langle\, g_2 \,\rangle$ as desired.

$\Leftarrow$ Suppose $g_1 \in \langle\, g_2 \,\rangle$. Then there exists an integer n so that $g_1 = g_2^n$. Now let $g \in \langle\, g_1 \,\rangle$ be arbitrary. This means there exists some integer m so that $g = g_1^m$. But then we have

$$g = g_1^m = (g_1)^m = (g_2^n)^m = g_2^{nm}.$$

In other words g is a power of g_2. Since g was an arbitrary element of $\langle g_1 \rangle$ we have proven that $\langle g_1 \rangle \subseteq \langle g_2 \rangle$ as desired. ■

LEMMA 5.77. *Suppose that $d \Big| |g|$. Then the order of g^d is $\frac{|g|}{d}$.*

PROOF. Let's denote the order of g as n, the integer (it's an integer since we are assuming that $d \mid n$) n/d as q, and the element g^d as h. With this notation, we are proving that $|h| = q$.

We have that

$$\begin{aligned} h^q &= (g^d)^q && (g^d = h) \\ &= g^{qd} \\ &= g^n && (n/d = q, \text{ i.e. } n = qd) \\ &= e_{\langle g \rangle}. && (|g| = n) \end{aligned}$$

Therefore the order of h is less than or equal to the integer q.

On the other hand, suppose $h^{q'} = e$ for some $q' \in \mathbb{Z}^+$. We need to prove that $q \leq q'$. But if we take both sides of the equation $h^{q'} = e_{\langle g \rangle}$ to the d^{th} power, we see

$$\begin{aligned} e_{\langle g \rangle} = h^{q'} &\Rightarrow e_{\langle g \rangle} = (g^d)^{q'} && (g^d = h) \\ &\Rightarrow e_{\langle g \rangle} = g^{q'd}. \end{aligned}$$

So, by the definition of what it means for $|g| = n$, $\langle n$ is least integer so that $g^n = e_{\langle g \rangle}$ we know that

$$n \leq q' \cdot d \Rightarrow \overset{q}{\overset{\shortparallel}{\frac{n}{d}}} \leq q' \Rightarrow q \leq q'$$

as desired. ■

LEMMA 5.78. *Let g be a group element of finite order. Then we have for any integer m*

$$\langle g^m \rangle = \left\langle g^{\gcd(m, |g|)} \right\rangle.$$

PROOF. Let's denote the order of g, by n and $\gcd(m, n)$ by d. We prove the lemma by showing that g^m is a power of g^d and vice versa, and then applying Lemma 5.76.

$\subseteq$ Then, since $d \mid m$ we have an integer $q \in \mathbb{Z}$, so that $m = qd$. But therefore

$$g^m = (g^{qd}) = (g^d)^q.$$

In other words g^m is a power of g^d.

$\supseteq$ By the division algorithm formula, there exists integers a and b so that

$$d = bm + an.$$

But then we have

$$\begin{aligned} g^d &= g^{bm+an} && (\text{Substitution}) \\ &= g^{bm} \cdot g^{an} \\ &= (g^m)^b \cdot (g^n)^a \\ &= (g^m)^b \cdot (e)^a && (\text{Since } |g| = n \text{ by def'n on } n) \\ &= (g^m)^b \cdot e && (\text{Id raised to power is id}) \\ &= (g^m)^b. \end{aligned}$$

■

NOTATION 5.79. In a group G denote

$$G\hat{}\,n \stackrel{\text{def}}{=} \{g^n : g \in G\}.$$

In other words, $G\hat{}\,n$ is the image of the function $G \to G$ which sends every element to its n^{th} power.

EXERCISE 5.58. In a cyclic group C of order n, prove that

SOLUTION

(1)

$$C\hat{}\,m = \langle x^m \rangle$$

where x is any generator of C.

(2)

$$C\hat{}\,m = C\left[\frac{n}{\gcd(n,m)}\right].$$

Hint! It suffices to prove this when $m \mid n$ where you need to prove

$$C\hat{}\,m = C\left[\frac{n}{m}\right].$$

Since, by the theorem above, the groups $\mathcal{U}(\mathbb{Z}_p)$ are cyclic of order $p-1$, this exercise implies that for an odd prime p, $\mathcal{U}(\mathbb{Z}_p)\hat{}\,2 = \mathcal{U}(\mathbb{Z}_p)\left[\frac{p-1}{2}\right]$. This means that $a \in \mathcal{U}(\mathbb{Z}_p)$ is a perfect square in $\mathcal{U}(\mathbb{Z}_p)$ iff $a^{\frac{p-1}{2}} \equiv_p 1$. Also, by a theorem below, this also means that half of all elements are perfect squares and half are not. We will come back to this later.

THEOREM 5.80. *For any integer m, the order of g^m is*

$$\frac{|g|}{\gcd(|g|, m)}.$$

PROOF. By the previous lemma, Lemma 5.78, the groups $\left\langle g^{\gcd(|g|,m)} \right\rangle$ and $\langle g^m \rangle$ are the same. Therefore since both elements generate exactly the same cyclic group, Theorem 5.61 says that both of the elements have the same order (since they each have the order equal to the order of this common group). Lastly, since $gcd(|g|, m)||g|$, two lemmas back, Lemma 5.77, says that

$$\left|g^{\gcd(|g|,m)}\right| = \frac{n}{\gcd(|g|, m)}$$

implying the same for g^m and we are done. ∎

Since in the group $\mathbb{Z}_n$, the element $[\,1\,]_n$ has order n, and since the element $[\,m\,]_n$ is $m \cdot [\,1\,]_n$, which is the "m^{th} power" in the additive group $\mathbb{Z}_n$, we have proven (we use the subscript + only to limit confusion between the order of this element and the multiplicative order of $[\,m\,]_n$ when thought of as an element of $\mathcal{U}(\mathbb{Z}_n)$)

COROLLARY 5.81.

$$|[\,m\,]_n|_+ = \frac{n}{\gcd(n, m)}.$$

∎

COROLLARY 5.82. *A cyclic group of n contains an element of order d iff $d \mid n$.*

Proof. Let C be a cyclic group and let x be a generator, i.e. an element of C of order n.

$\Rightarrow$ Every element of C is a power of x so by Exercise 5.26 we know that the order of every element divides $|x| = n$.

$\Leftarrow$ If $d \mid n$, then $x^{n/d}$ is an element of order d in C. ■

We could have also proved the $\Rightarrow$ direction simply by using the formula to note that the order of every element of C has the form

$$\frac{n}{\gcd(n,m)},$$

which as $\gcd(n,m) \mid n$, is a divisor of n. Actually, we will eventually prove something much stronger than this as a corollary of Lagrange's theorem. Namely, in any group, the order of any element divides the order of the group. Since the order of $\langle g \rangle$ is $|g|$, this theorem applies to this case.

Corollary 5.83. *g^m generates $\langle g \rangle$ iff m and $|g|$ are relatively prime.* ■

Corollary[2] 5.84. *There are $\vartheta(n)$ generators of a cyclic group of order n.* ■

Here we assume that p is an odd prime number.

Definition 5.85. *$\mathbb{Z}_p$ contains an i* if there exists an $x \in \mathbb{Z}_p$ so that $x^2 \equiv_p -1$.

For instance $\mathbb{Z}_{13}$ contains an i since $5^2 = 25 \equiv_{13} -1 \bmod 13$. Also $\mathbb{Z}_{29}$ contains an i since $12^2 \equiv_{29} -1$.

Example /Theorem 5.86. $\mathbb{Z}_p$ contains an i iff $p \equiv_4 1$.

Proof. $\mathbb{Z}_p$ contains an i iff there there exists an x with $x^2 \equiv_p -1$. But since we are assuming that p is odd, $-1 \neq 1$ and -1 is the only element of $\mathcal{U}(\mathbb{Z}_p)$ of order 2. Hence $\mathbb{Z}_p$ contains an i iff $\mathcal{U}(\mathbb{Z}_p)$ contains an element of order 4. But since $\mathcal{U}(\mathbb{Z}_p)$ is cyclic by Theorem 5.75, we can apply Corollary 5.82 to conclude that this happens iff 4 divides p, i.e. iff $p \equiv_4 1$. ■

For instance we have that $\mathbb{Z}_p$ contains an i for $p = 5, 13, 17, 29, 37$ but not for $p = 3, 7, 11, 19, 23$, etc.

Example 5.87. The order of $[\,18\,]_{30}$ is 5. This is because $[\,18\,]_{30} = 18 \cdot [\,1\,]_{30}$, the order of $[\,1\,]_{30}$ is 30, and that the gcd of 30 and 18 is 6 while 30/6=5. To make us feel better let's verify this directly:

$$2 \cdot 18 = 36 \equiv_{30} 6 \qquad 3 \cdot 18 \equiv_{30} 6 + 18 = 24 \qquad 4 \cdot 18 \equiv_{30} 2 \cdot 6 = 12 \qquad 5 \cdot 18 = 90 \equiv_{30} 0.$$

Notice that 6 shows up as a power (multiple) of 18 as guaranteed by the proof of the theorem.

One should notice that the two lemmas are special cases of the theorem. This is because if $m \mid n$, then $\gcd(m,n) = m$ and if instead they were relatively prime, then their gcd would be 1.

Corollary 5.88. *A cyclic group of even order has exactly one element of order 2. A cyclic group of odd order contains none.*

Proof. Let n denote the order of g. If 2 divides n, then $g^{\frac{n}{2}}$ is an element of order 2. On the other hand, if $n/\gcd(n,m) = 2$, then $\gcd(n,m) = n/2$. But the only divisors of n with this property is $m = n/2$. As such $g^{\frac{n}{2}}$ is the only element of order 2 in a cyclic group. ■

For C cyclic, $\mathbf{O}_2(C) \leq$

Recall that $\vartheta_n(m)$ is defined to be the number of positive integers less than m which are relatively prime to m whenever m divides n and 0 otherwise. We can greatly expand the previous corollary, Corollary 5.88 as follows.

COROLLARY 5.89. *Suppose C is a cyclic group of order n. Then we have for any integer m*

$$\mathbf{o}_m C = \vartheta_n(m) \quad \textit{and} \quad \mathbf{O}_m C = \gcd(m, n).$$

REMARK 5.90. In more detail, this means that if $m \mid n$, then there are exactly m solutions to the equation $x^m = e$ in the group C.

PROOF. Since we have that every element of $g \in C$, $g^n = e_C$, $C[n] = C$. Therefore, for any integer m,

$$C[m] = C[m] \cap C = C[m] \cap C[n] = C[\gcd(m, n)]$$

where we've used Exercise 5.27 back on page 265.

Moreover, by Exercise 5.58, we know that for any generator x of C,

$$\left\langle x^{\frac{n}{\gcd(n,m)}} \right\rangle = C\hat{}\,\tfrac{n}{\gcd(n,m)} = C[\gcd(n, m)] = C[m].$$

But since $x^{\frac{n}{\gcd(n,m)}}$ has order $\gcd(n, m)$ so does $C[m]$, proving $\mathbf{O}_m C$ – which by definition is $|C[m]|$ – is equal to $\gcd(m, n)$.

To prove the first equality, note if $g \in C$ has order m, then $m \mid n$, by Corollary 5.82. Meanwhile if $m \nmid n$, then $\vartheta_n(m) = 0$. The first equality checks as $0 = 0$. So we need to prove the first equality in the special case that $m \mid n$. In this case, we have $\vartheta_n(m) = \vartheta(m)$. Now by the argument of the first part of the proof, this means that g is a generator of the cyclic group $\left\langle x^{\frac{n}{m}} \right\rangle$ which has order m. But Corollary 5.84 back on page 293 says that a cyclic group of order m has $\vartheta(m)$ generators. ∎

EXAMPLE 5.91.

(1) The group $\mathcal{U}(\mathbb{Z}_{36})$ is not cyclic. The fastest way of proving this is by noting that

$$17^2 = 289 = 8 \cdot 36 + 1,$$

and so $[\,17\,]_{36}$ has (multiplicative) order 2. Therefore, $[\,35\,]_{36}$ and $[\,17\,]_{36}$ give us distinct elements of order 2 in $\mathcal{U}(\mathbb{Z}_{36})$ and hence this group ($\mathcal{U}(\mathbb{Z}_{36})$) cannot be cyclic.

(2) The group $\mathcal{U}(\mathbb{Z}_{35})$ is not cyclic since $[\,6\,]_{35}$ and $[\,34\,]_{35}$ each have order 2 in this group.

EXERCISE 5.59. Let r be an integer greater than 2. Prove that the group $\mathcal{U}(\mathbb{Z}_{2^r})$ is not cyclic.
Hint! Show that $[\,-1\,]_{2^r}$ and $[\,2^{r-1} - 1\,]_{2^r}$ are distinct elements of order 2. (Be sure to note exactly where you used the fact that $r > 2$!)
SOLUTION

EXERCISE 5.60. Generalize the example 5.91**(2)** to prove that for distinct odd primes p and q, that the abelian group $\mathcal{U}(\mathbb{Z}_{p \cdot q})$ is not cyclic.
Hint! By the Chinese Remainder Theorem, there exists an integer $m \in \mathbb{Z}$ so that $m \equiv_p 1$ and $m \equiv_q -1$. Prove that we have that $[\,m\,]_{pq} \neq [\,-1\,]_{pq}$ and that $[\,m\,]_{pq}$ it has (multiplicative) order 2.
SOLUTION

5.9.4 Summary About Orders

So let's collect all the results about orders we have proven so far (found in the exercises, theorems, and lemmas in the last section).

(1) Every integer, rational number, real number, complex number, and quaternion (except for 0) has infinite (additive) order.

(2) The (additive) order of $[\,a\,]_n \in \mathbb{Z}_n$ is calculated using the formula

$$|\,[\,a\,]_n\,| = \frac{n}{\gcd(a,n)}.$$

More generally, in any cyclic group, $\langle\, g\,\rangle$ we have that

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

For instance, in $\mathcal{D}_{20}$, we have that a has order 20 and ρ^8 has order 5.

(3) Let

$$\mathsf{q} = q_R + \mathbf{q}_P$$

be a quaternion.

- If q is not in $\mathbb{S}^3$ (i.e. if $q_R^2 + \|\,\mathbf{q}_P\,\|^2 \neq 1$) then q has infinite order.
- If q is in $\mathbb{S}^3$ (i.e. if $q_R^2 + \|\,\mathbf{q}_P\,\|^2 = 1$) then the real part, q_R, is between -1 and 1 and if arccos of this number has the form

$$\frac{m}{n} \cdot (2\pi)$$

with $\frac{m}{n}$ a reduced fraction (i.e. with m and n relatively prime), then the order of q is n. (This question therefore depends on one's ability to recognize a decimal as a rational number, which in general, is impossible.)

(4) For a matrix in $\mathbf{SL}_2(\mathbb{Z})$ there are several cases based on the trace.

- The order of identity is 1.
- The order of negative the identity,

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

is 2. This is the only matrix whose order is 2.

- If the trace is 0 (and determinant 1), then the order is 4. For instance, the matrices

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 3 & -2 \\ 5 & -3 \end{pmatrix}$$

have order 4. Also, every matrix of the form

$$\left\{ \begin{pmatrix} n & -1 \\ n^2 - 1 & -n \end{pmatrix} : n \in \mathbb{Z} \right\}$$

has order 4.

- If the trace is 1 (and determinant 1), then the order is 6. For instance the matrices

$$\begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -4 & 7 \\ -3 & 5 \end{pmatrix}$$

  have order 6.
- If the trace is -1 (and determinant 1), then the order is 3. For instance, the matrix

$$\begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$$

  has order 3.
- If the trace is 2 (and determinant 1) and is not the identity, the matrix has infinite order. For instance the matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -3 & 4 \\ -4 & 5 \end{pmatrix}$$

  have infinite order.
- All other matrices in $\mathbf{SL}_2(\mathbb{Z})$ have infinite order. For instance the matrices

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 5 & 7 \\ 2 & 3 \end{pmatrix}$$

  have infinite order.

5.9.5 Using Möbius Inversion*

In the Appendix, K.K.6, we discuss the Möbius inversion formula. Here's a result that uses facts (and notation) from that appendix. In fact, the non-trivial part is Example K.14 there.

Theorem 5.92.

$$\mathbf{O}_n(G) = \sum_{d|n} \mathbf{o}_n(G) \quad \textit{and} \quad \mathbf{o}_n(G) = \sum_{d|n} \mu(d) \cdot \mathbf{O}_{\frac{n}{d}}(G).^{\text{XVIII}} \tag{5}$$

Proof. The first sum is nothing more than a translation of Theorem 5.38**(2)**. The second is due to a general number theory fact discussed in Appendix K. ∎

Example 5.93. We have

$$\mathbf{o}_{60}(G) = \mathbf{O}_{60}(G) - (\mathbf{O}_{30}(G) + \mathbf{O}_{20}(G) + \mathbf{O}_{12}(G)) + (\mathbf{O}_4(G) + \mathbf{O}_6(G) + \mathbf{O}_{10}(G)) - 1.$$

We can use this to compute the number of elements of any given order in $H \times K$ when we can compute those things in H and K. Now, in general, an element $g = (h, k) \in H \times K$ we have that $g^n = (h^n, k^n)$. So $g^n = e_{H\times K}$ iff $h^n = e_H$ and $k^n = e_K$. We have just proven:

[XVIII] The function μ is defined in the same appendix, Appendix K.

PROPOSITION 5.94. *For any integer m*

$$\mathbf{O}_m(H \times K) = \mathbf{O}_m(H) \cdot \mathbf{O}_m(K).$$

■

COROLLARY 5.95. *For any integer n,*

$$n = \sum_{d|n} \vartheta(d) \quad \text{and} \quad \frac{\vartheta(n)}{n} = \sum_{d|n} \frac{\mu(d)}{d}. \tag{6}$$

PROOF. Corollary 5.89 says that in the cyclic group $\mathbb{Z}_n$ we have for any $d \mid n$, $\mathbf{o}_d(\mathbb{Z}_n) = \vartheta(d)$ and $\mathbf{O}_d(\mathbb{Z}_n) = d$. Plugging these into the previous theorem gives the result. ■

Of course, giving an equation for something we understand poorly in terms of something we understand well vs giving us something we understand well in terms of something we understand poorly, the right hand side of (6) is more interesting. For instance we have

COROLLARY 5.96.

(1)

$$\frac{\vartheta(n)}{n} = 1 + \sum_{p_1 \in P} \frac{(-1)}{p_1} + \sum_{\{p_1,p_2\} \in \mathcal{P}_2(P)} \frac{1}{p_1 p_2} + \cdots + \sum_{\{p_1,\ldots,p_i\} \in \mathcal{P}_i(P)} \frac{(-1)^i}{p_1 \cdots p_i} + \cdots$$

where P is the set of prime divisors of n.

(2)

$$\frac{\vartheta(n)}{n} = \prod_{\substack{p|n \\ p>0 \text{ prime}}} \left(1 - \frac{1}{p}\right).$$

PROOF. We get the first part by explicitly plugging $\mu(d)$ into (6) and the second part is a simplification of the first as can easily be seen by expanding the second. ■

For instance, if $n = p^2 \cdot q \cdot r$, with p, q, r distinct primes, the first sum would read

$$\frac{\vartheta(n)}{n} = 1 - \left(\frac{1}{p} + \frac{1}{q} + \frac{1}{r}\right) + \left(\frac{1}{pq} + \frac{1}{pr} + \frac{1}{rq}\right) - \left(\frac{1}{pqr}\right)$$

while the second product would read

$$\frac{\vartheta(n)}{n} = \left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{q}\right)\left(1 - \frac{1}{r}\right).$$

Note that these ratios care less about the second p which divides n.

5.10 CONJUGATION

DEFINITION 5.97. Let $a, b \in G$. Then a is *conjugate* (in G) to b via $g \in G$ if $g\,a\,g^{-1} = b$. Generically, we call a and b conjugate if they are conjugate via some element.

While in a group $ab \neq ba$ these two elements are always conjugate. Indeed $b(ab)b^{-1} = ba$. We will see that conjugate elements have mostly the same properties. For instance, conjugate elements have the same order. The result for $G = \mathcal{S}_n$ is particularly nice, see Theorem 7.40.

WARNING! 5.98. This has nothing to do with what the conjugate of a complex number or a quaternion! Note in that case, the conjugate of an element makes sense by itself (so it's a uniary operation), here, to conjugate in a group, you need two elements (so it's a binary operation)!

DEFINITION 5.99. Define a relation $\sim_G$ on a group G by declaring

"$a \sim_G b$ iff a is conjugate to b".

That is $a \sim_G b$ iff there exists a $g \in G$ so that $g\,a\,g^{-1} = b$.

THEOREM 5.100. *The relation $\sim_G$ is an equivalence relation.*

PROOF.

Reflexive Let $g \in G$. Then we have $e_G g e_G^{-1} = g$. Therefore $g \sim_G g$ and $\sim_G$ is reflexive.

Symmetric Suppose $a \sim_G b$. Then, there exists $g \in G$ so that $g\,a\,g^{-1} = b$. But, since G is a group, $g^{-1} \in G$. Moreover, $g^{-1}\,b\,g = g^{-1}gag^{-1}g = a$. Therefore $b \sim_G a$ and $\sim_G$ is symmetric.

Transitive Suppose that $a \sim_G b$ and $b \sim_G c$ for elements $a, b, c \in G$. Then there exists elements $g_1, g_2 \in G$ so that

$$g_1 a g_1^{-1} = b \quad \text{and} \quad g_2 b g_2^{-1} = c.$$

Therefore

$$(g_2\,g_1)\,a\,(g_2\,g_1)^{-1} = g_2\,(g_1\,a\,g_1^{-1})\,g_2^{-1} = g_2\,b\,g_2^{-1} = c.$$

Therefore $a \sim_G c$ and $\sim_G$ is transitive. ■

We will present a much slicker proof of the following exercise in Theorem 8.35 coming up on page 401.

EXERCISE 5.61.

SOLUTION

(1) Prove for any $g, a \in G$ and any integer n we have that

$$g\,a^n\,g^{-1} = \left(gag^{-1}\right)^n.$$

(2) Use the first part to prove that $a \sim_G b \Rightarrow a^n \sim_G b^n$. [1]

(3) Use the first part to prove that if $a \sim_G b \Rightarrow |a| = |b|$.

[1] In words, if a and b are conjugate, then so is any power of a and b. This includes negative powers, e.g.

$$a \sim_G b \Rightarrow a^{-1} \sim_G b^{-1}.$$

EXERCISE 5.62.

(1) Find a non-abelian group where every element is conjugate to its inverse.

(2) Find a non-abelian group where not every element is conjugate to its inverse.
Hint! $\mathcal{D}_n$, $\mathcal{S}_n$ never work for this second case.

SOLUTION

EXERCISE 5.63. Suppose that H is a subgroup of G.

(1) Prove conjugacy in H implies conjugacy in G. That show for all $h_1, h_2 \in H$

$$h_1 \sim_H h_2 \Rightarrow h_1 \sim_G h_2.$$

(2) Prove that the converse is false (of course this is done by finding an explicit example!)

SOLUTION

EXERCISE 5.64. Suppose $a \sim a^{-1}$ via g. Prove that g^2 commutes with a but g commutes with a iff $a^2 = e$.

SOLUTION

DEFINITION 5.101. An equivalence class of some $a \in G$ under the relation $\sim_G$ is called the *conjugacy class* of a in G. It is denoted as $\mathfrak{cc}_G(a)$.

Therefore, in less fancy language the conjugacy class of an element is the set of elements which are conjugate to the given element, i.e.

$$\mathfrak{cc}_G(a) = \{b \in G : \exists g \in G \text{ s.t. } gag^{-1} = b\} \quad \text{or} \quad = \{gag^{-1} : g \in G\}.$$

EXAMPLE /THEOREM 5.102. Two elements of $\mathbb{S}^3$ are conjugate iff their real parts are equal.

PROOF.

$\Leftarrow$ Let q and q' be two quaternions whose real parts are equal. Then there exists an angle θ and $\widehat{\mathbf{u}}, \widehat{\mathbf{u}}' \in \mathbb{S}^2$ so that

$$\mathsf{q} = \cos\theta + \sin\theta\,\widehat{\mathbf{u}} \quad \text{and} \quad \mathsf{q}' = \cos\theta + \sin\theta\,\widehat{\mathbf{u}}'.$$

Then I claim that

$$\left[\frac{\sqrt{2}}{2}(\widehat{\mathbf{u}} + \widehat{\mathbf{u}}')\right] \widehat{\mathbf{u}} \left[\frac{\sqrt{2}}{2}(\widehat{\mathbf{u}} + \widehat{\mathbf{u}}')\right]^{-1} = \widehat{\mathbf{u}}'.$$ II

Indeed, cross multiplying and canceling the scalar $\frac{\sqrt{2}}{2}$ from both sides, we see that it suffices to show that

$$(\widehat{\mathbf{u}} + \widehat{\mathbf{u}}')\widehat{\mathbf{u}} = \widehat{\mathbf{u}}'(\widehat{\mathbf{u}} + \widehat{\mathbf{u}}').$$

But this is straight forward, indeed

$$\begin{aligned}(\widehat{\mathbf{u}} + \widehat{\mathbf{u}}')\widehat{\mathbf{u}} &= \widehat{\mathbf{u}}\widehat{\mathbf{u}} + \widehat{\mathbf{u}}'\widehat{\mathbf{u}} \\ &= -1 + \widehat{\mathbf{u}}'\widehat{\mathbf{u}}\end{aligned}$$

while

$$\begin{aligned}\widehat{\mathbf{u}}'(\widehat{\mathbf{u}} + \widehat{\mathbf{u}}') &= \widehat{\mathbf{u}}'\widehat{\mathbf{u}} + \widehat{\mathbf{u}}'\widehat{\mathbf{u}}' \\ &= \widehat{\mathbf{u}}'\widehat{\mathbf{u}} - 1 \\ &= -1 + \widehat{\mathbf{u}}'\widehat{\mathbf{u}}\end{aligned}$$

and the two are equal as promised. Therefore, we have

$$\begin{aligned}(\widehat{\mathbf{u}} + \widehat{\mathbf{u}}')\mathsf{q}(\widehat{\mathbf{u}} + \widehat{\mathbf{u}}')^{-1} &= (\widehat{\mathbf{u}} + \widehat{\mathbf{u}}')\cos\theta(\widehat{\mathbf{u}} + \widehat{\mathbf{u}}')^{-1} + \sin\theta(\widehat{\mathbf{u}} + \widehat{\mathbf{u}}')\widehat{\mathbf{u}}(\widehat{\mathbf{u}} + \widehat{\mathbf{u}}')^{-1} \\ &= \cos\theta + \sin\theta\,\widehat{\mathbf{u}}' \\ &= \mathsf{q}'.\end{aligned}$$

and q and q' are conjugate.

⇒ In the appendix, we prove that the conjugate of a pure quaternion is a pure quaternion (this is not difficult and is straight forward). Thus for every q, $a + b\widehat{\mathbf{u}} \in \mathbb{S}^3$ we have

$$\mathsf{q}(a + b\widehat{\mathbf{u}})\mathsf{q}^{-1} = a + b\widehat{\mathbf{u}}'$$

where $\widehat{\mathbf{u}}'$ is pure. Hence the real part of $a + b\widehat{\mathbf{u}}$ and its conjugate by q are both a. ■

EXERCISE 5.65. Define a binary operation on a group G via $g \circ h := ghg^{-1}$.

SOLUTION

(1) Prove that $(g \circ h) \circ k = g \circ (h \circ k)$ iff g and k commute.

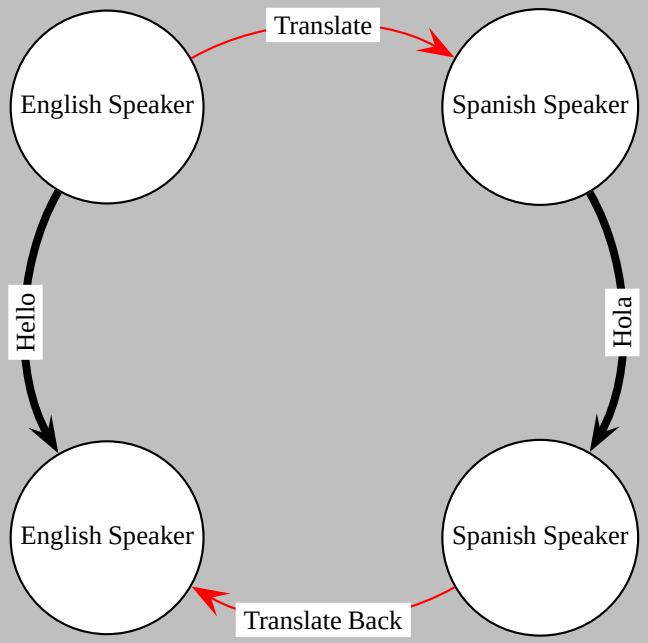
(2) Prove that TFAE

- The operation $\circ$ is trivial;
- The operation $\circ$ is associative;
- The group G is abelian.

[II] The $\frac{\sqrt{2}}{2}$ factor is there just so everything happens in $\mathbb{S}^3$.

RANT 5.103. It surprises most students that conjugation shows up as much as a it does in mathematics. In fact, in many groups including for quaternions, permutations, and of course abelian groups, conjugation is actually easier to compute than multiplication! [III]

Consider the following story which, I hope, will make conjugation seem more natural. Two people, one who speaks English but not Spanish the other in the opposite situation: speaks Spanish but not English, wish to communicate. With this goal in mind they hire a translator. One way to think of this translation is as



With this in mind it's useful to think of conjugation in terms of a translation.[IV]

EXERCISE 5.66. Perform the calculation

$$\text{CCCXCIV} + \text{MMMDXLIX} = ?.$$

The numbers are given in roman numerals. Give your answer in terms of roman numerals. Explain the relevance to Rant 5.103.

SOLUTION

5.11 EXAMPLES

The examples in this section will be discussed over and over and over and over and ... yeah.

EXAMPLE 5.104.

(1) The easiest groups to deal with are the groups $(\mathbb{Z}_n, +)$.[I] We include in this the case that $n = 0$ in which case $\mathbb{Z}_0 = \mathbb{Z}$. While these groups are very simple, it turns out that every abelian group is a direct product of these groups.

(2) We have already seen that the groups $\mathcal{U}(\mathbb{Z}_n)$ are a bit more difficult to get a handle on. For instance figuring out which element is the inverse of a given element can be quite a challenge. The table for $\mathcal{U}(\mathbb{Z}_{14})$ is as follows:

·	1	3	5	9	11	13
1	1	3	5	9	11	13
3	3	9	1	13	5	11
5	5	1	11	3	13	9
9	9	13	3	11	1	5
11	11	5	13	1	9	3
13	13	11	9	5	3	1

REMARK 5.105. We can save a lot of time in writing out the multiplication table by using the symmetry present in certain groups. In the present situation, if we write half of the elements of $\mathcal{U}(\mathbb{Z}_n)$ as negatives of the first half, then we only need to figure out $1/8^{\text{th}}$ of the multiplication table! This is because if we flip the table across the middle along the tops or along the bottom, we simply switch all of the entries to their negatives. For instance, if we flip along the top middle we are replacing $a \cdot b$ with $a \cdot (-b) = -a \cdot b$. If, on the other hand, we flip along the diagonal, we are replacing $a \cdot b$ with $b \cdot a$. But since $\mathcal{U}(\mathbb{Z}_n)$ is always abelian these two are the same. For instance, in the example we just did with $n = 14$

·	1	3	5	−5	−3	−1
1	1	3	5	−5	−3	−1
3	3	−5	1	−1	5	−3
5	5	1	−3	3	−1	−5
−5	−5	−1	3	−3	1	5
−3	−3	5	−1	1	−5	3
−1	−1	−3	−5	5	3	1

EXERCISE 5.67. Make a multiplication table for $\mathcal{U}(\mathbb{Z}_{30})$. (While using the trick just illustrated you only need to make 6 calculations as opposed to the 49 you would have to do otherwise (not counting calculations with identity as calculations!).

SOLUTION

(3) The second easiest class of groups to work with are probably the dihedral groups, $\mathcal{D}_n$, discussed earlier. As we already agreed, if we wanted to talk about a specific n-gon we typically talk about the n-gon whose vertices are $\{\{n\}\}$, but for most purposes it's good enough to think of an abstract n-gon, for instance one you cut out of

[I] Notice that $(\mathbb{Z}_n, \cdot)$ is never a group since the element $[\,0\,]_n$ can never have a multiplicative inverse.

paper. For consistancy, we can define a 2-gon as a rectangle which is not a square (this makes the statement of some results a bit easier). The basic rotation, by $\frac{2\pi}{n}$ will always be denoted by ρ, and a reflection (which in $\{\{n\}\} \subseteq \mathbb{C}$ corresponds to complex conjugation) with always be denoted by $\mathbf{r}$. Note that ρ has a different meaning for different integers n. Using complex numbers and recalling the notation

$$\chi_n \stackrel{\text{def}}{=} \exp\left(\frac{2\pi}{n} i\right) = \cos\frac{2\pi}{n} + \sin\frac{2\pi}{n}\, i.$$

we have that $\rho(z) = \chi_n(z)$ and so

$$\begin{aligned}
\mathbf{r}(\rho(\mathbf{r}^{-1}(z))) &= \mathbf{r}(\rho(\overline{z})) \\
&= \mathbf{r}(\chi_n \overline{z}) \\
&= \overline{\chi_n \overline{z}} \\
&= \overline{\chi_n}\, \overline{\overline{z}} \\
&= \chi_n^{-1} z \\
&= \rho^{-1}(z).
\end{aligned}$$

This implies that $\mathbf{r} \circ \rho \circ \mathbf{r}^{-1}$ is always equal to ρ^{-1}. In fact, using basic planar geometry, see for instance Corollary C.40 coming up on page 782 in the appendix, one can show for any reflection R and rotation ϱ of the plane we always have

$$R \circ \varrho \circ R^{-1} = \varrho^{-1}.$$

One way of thinking of this is by saying that every time you pass a ρ by a $\mathbf{r}$ you change ρ to a ρ^{-1}. For instance, no matter what planar rotation and reflection we are dealing with,

$$\begin{aligned}
\rho^3 \mathbf{r} \rho^5 \mathbf{r} \rho^2 &= \rho^3 \rho^{-5} \mathbf{r}\mathbf{r} \rho^2 \\
&= \rho^3 \rho^{-5} \rho^2 \\
&= \rho^0 \\
&= \mathbf{1}.
\end{aligned}$$

THEOREM 5.106. *Every element of $\mathcal{D}_n$ can be written uniquely in the form*

$$\mathbf{r}^\epsilon \rho^i$$

where $\epsilon \in \{0, 1\}$ and $i \in \{0, \ldots, n-1\}$.

PROOF. We need to prove two things:

(a) (Existence) Suppose that $\varphi \in \mathcal{D}_n$. Then φ applied to the complex number $1 = 1 + 0i$ is some vertex of the regular n–gon $\{\{n\}\}$. But every such vertex is equal to $\chi_n^i = \exp(2\pi i/n)$ for some integer $0 \leq i < n$. Thus

$$\rho^{-i}(\varphi(1)) = \rho^{-i}(\chi_n^i) = \chi_n^{-i} \cdot \chi_n^i = 1.$$

Now since φ is a symmetry, and χ_n is adjacent to 1, $\varphi(\chi_n)$ must be adjacent to $\varphi(1) = \chi_n^i$. This means that $\varphi(\chi_n)$ must be χ_n^{i-1} or χ_n^{i+1}. In the latter case we have that $\rho^{-i}\varphi(\chi_n) = \chi_n$ and in the former case we have that $\mathbf{r}\rho^{-i}(\varphi(\chi_n)) = \chi_n$. In the latter case we have that $\mathbf{r}\rho^{-i}\varphi$ fixes two vertices 1 and χ_n. This clearly implies

that $\mathbf{r}\rho^{-i}\varphi = \mathbf{1}$. But this means that

$$\begin{aligned} \mathbf{r}\rho^{-i}\varphi &= \mathbf{r} \\ \Rightarrow \rho^{-i}\varphi &= \mathbf{r} \\ \Rightarrow \varphi &= \rho^{i}\mathbf{r}. \end{aligned}$$

Similarly, in the former case we can see that $\varphi = \rho^i$. Since φ was arbitrary we have proven that every element of $\mathcal{D}_n$ can be written in the stated form.

(b) (Uniqueness) The element $\rho^i\mathbf{r}^\epsilon$ takes the vertex 1 to χ_n^i and χ_n to $\chi_n^{i+(-1)^\epsilon}$. Hence if

$$\rho^i\mathbf{r}^\epsilon = \rho^{i'}\mathbf{r}^\epsilon$$

then

$$\chi_n^i = \chi_n^{i'} \tag{7}$$

$$\chi_n^{i+(-1)^\epsilon} = \chi_n^{i'+(-1)^{\epsilon'}}. \tag{8}$$

And since $0 \leq i, i' < n$ and χ_n has order n, (7) implies that $i = i'$. Now using this and (8) we see that we also have that $\epsilon = \epsilon'$, as desired. ■

COROLLARY 5.107. $|\mathcal{D}_n| = 2n$.

PROOF. We have just seen that the elements of $\mathcal{D}_n$ are

$$\{\mathbf{1}, \rho, \rho^2, \ldots, \rho^{n-1}, \mathbf{r}, \rho\mathbf{r}, \rho^2\mathbf{r}, \ldots, \rho^{n-1}\mathbf{r}\}$$

with no repeats. There are $2n$ elements in this set so there are $2n$ elements in $\mathcal{D}_n$. ■

EXAMPLE / THEOREM 5.108. $\mathcal{D}_n$ contains an element of order m iff $m = 2$ or $m \mid n$. Moreover

$$\mathbf{o}_m(\mathcal{D}_n) = \begin{cases} \vartheta(m) \text{ if } m \mid n \text{ and } m \neq 2 \\ n+1 \text{ if } m = 2 \text{ and } n \text{ is even.} \\ n \text{ if } m = 2 \text{ and } n \text{ is odd.} \end{cases}$$

where ϑ is the Euler totient function.

PROOF. Each of the n elements $\rho^i\mathbf{r}$ are reflections and hence have order 2. Meanwhile each of the elements ρ^i are in the cyclic group $\langle \rho \rangle$ which has order $|\rho| = n$. By a previous result, if $m \mid n$, then there are $\vartheta(n)$ elements of order n in $\langle \rho \rangle$ and none in there otherwise. Putting this all together gives the result. ■

(4) The third easiest class of class of groups (often called *generalized quaternion groups*), are denoted $B\mathcal{D}_n$ and form a group under the operation of quaternion multiplication. In this text, we will call these groups *double dihedral groups* or (if n is fixed) the *binary n-gon group* (in the literature, they are typically called generalized quaternion groups, but we will be rebels and not call them that .. I hope the reason why will become clear as well go on). Now for the definition. As a set (we fix the integer d instead of n and reserve n to be a dummy index) it is

$$B\mathcal{D}_d \stackrel{\text{def}}{=} \{\mathbf{j}^n\chi_{2d}^m : n, m \in \mathbb{Z}\}. \tag{9}$$

Here, as defined in Definition 3.59 back on page 98, χ_d is the $2d^{\text{th}}$ root of unity

$$e^{\frac{\pi i}{d}} = \cos\frac{\pi}{d} + \sin\frac{\pi}{d}\,\mathbf{i}.$$

If d is understood or irrelevant we drop the d and simply denote this as χ. Since $\mathbf{j}$ has (multiplicative) order 4, we needn't consider integers n in the definition besides those between 0 and 3, i.e. $0 \le n < 4$ in the above. At the same time, since

$$\chi^d = e^{\pi i} = -1$$

and for $m = d + m'$

$$\chi^m = \chi^d\chi^{m'} = -\chi^{m'} = \mathbf{j}^2\chi^{m'}.$$

Therefore, since we get -1 from a power of $\mathbf{j}$, we only need to consider power of χ between 0 and $d-1$ in the definition. Therefore we can simplify the original definition given in (9) down to

$$B\mathcal{D}_d = \{\mathbf{j}^n\chi^m : 0 \le n < 4, 0 \le m < d\}.$$

In this way we can write every element of $B\mathcal{D}_d$ as

$$\begin{array}{rlllll}
1 & \chi & \chi^2 & \cdots & \chi^{d-1} & (10)\\
-1 & -\chi & -\chi^2 & \cdots & -\chi^{d-1} & (11)\\
\mathbf{j} & \mathbf{j}\chi & \mathbf{j}\chi^2 & \cdots & \mathbf{j}\chi^{d-1} & (12)\\
-\mathbf{j} & -\mathbf{j}\chi & -\mathbf{j}\chi^2 & \cdots & -\mathbf{j}\chi^{d-1}. & (13)
\end{array}$$

PROPOSITION 5.109. *There are no repeats in this list.*

PROOF. Indeed, no element in the first two rows, (10) and (11), can repeat since $|\chi| = 2d$ and these first two rows make up $\langle\chi\rangle$. At the same time since every element in the last two rows, (12) and (13), since each element in those rows is of the form j times a unique element in one of the first two rows. Lastly, no element in one of the first two rows can equal an element in one of the last two rows since no element in the first two rows in pure while every element in the last two rows is pure. ■

Therefore

COROLLARY 5.110.

$$|B\mathcal{D}_n| = 4n.$$

■

Other sexy ways or writing elements of $B\mathcal{D}_d$ include

$$B\mathcal{D}_d = \{(\pm 1)\mathbf{j}^n\chi^m : 0 \le n < 2, 0 \le m < d\}$$

or

$$B\mathcal{D}_d = \{\mathbf{j}^n\chi^m : 0 \le n < 2, 0 \le m < 2d\}.$$

Of particular importance is the case that $d = 2$. There $\chi_{2\cdot 2} = \mathbf{i}$, and the $8 = 4\cdot 2$ elements are

$$B\mathcal{D}_2 = \{\pm 1, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}\} = \{1, -1, \mathbf{i}, -\mathbf{i}, \mathbf{j}, -\mathbf{j}, \mathbf{k}, -\mathbf{k}\}.$$

We will follow the usual convention and denote this group as Q_8 instead of $B\mathcal{D}_2$. This group is called the *Quaternion Group* (instead of the binary rectangle group). Beware, this is different from the set $\mathbb{H}^\times$. I will typically call this group the *Quaternion group of order 8*. Here is a multiplication table for Q_8.

Q_8	1	$\mathbf{i}$	-1	$-\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$	$-\mathbf{j}$	$-\mathbf{k}$
1	1	$\mathbf{i}$	-1	$-\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$	$-\mathbf{j}$	$-\mathbf{k}$
$\mathbf{i}$	$\mathbf{i}$	-1	$-\mathbf{i}$	1	$\mathbf{k}$	$-\mathbf{j}$	$-\mathbf{k}$	$\mathbf{j}$
-1	-1	$-\mathbf{i}$	1	$\mathbf{i}$	$-\mathbf{j}$	$-\mathbf{k}$	$\mathbf{j}$	$\mathbf{k}$
$-\mathbf{i}$	$-\mathbf{i}$	1	$\mathbf{i}$	-1	$-\mathbf{k}$	$\mathbf{j}$	$\mathbf{k}$	$-\mathbf{j}$
$\mathbf{j}$	$\mathbf{j}$	$-\mathbf{k}$	$-\mathbf{j}$	$\mathbf{k}$	-1	$\mathbf{i}$	1	$-\mathbf{i}$
$\mathbf{k}$	$\mathbf{k}$	$\mathbf{j}$	$-\mathbf{k}$	$-\mathbf{j}$	$-\mathbf{i}$	-1	$\mathbf{i}$	1
$-\mathbf{j}$	$-\mathbf{j}$	$\mathbf{k}$	$\mathbf{j}$	$-\mathbf{k}$	1	$-\mathbf{i}$	-1	$\mathbf{i}$
$-\mathbf{k}$	$-\mathbf{k}$	$-\mathbf{j}$	$\mathbf{k}$	$\mathbf{j}$	$\mathbf{i}$	1	$-\mathbf{i}$	-1

The next easiest case (since 3 is the next number after 2) to look at is $B\mathcal{D}_3$. The $4 \cdot 3 = 12$ elements of $B\mathcal{D}_3$ are

$$\{1, \chi, \chi^2, -1, -\chi, -\chi^2, \mathbf{j}, \mathbf{j}\chi, \mathbf{j}\chi^2, -\mathbf{j}, -\mathbf{j}\chi, -\mathbf{j}\chi^2\} \tag{14}$$

where, in this case, χ is 6^{th} root of unity:

$$\chi = \exp\frac{\pi i}{3} = \cos\frac{\pi}{3} + \cos\frac{\pi}{3}i = \frac{1}{2} + \frac{\sqrt{3}}{2}i = \frac{1}{2} + \frac{\sqrt{3}}{2}\mathbf{i}..$$

Lastly we show that these $B\mathcal{D}_d$ dudes are always groups. By Theorem 5.53 we simply need to show that $B\mathcal{D}_d$ is closed under (quaternion) multiplication. Indeed, for integers $n, m, \dot{n}$ and $\dot{m}$ we have that

$$\begin{aligned}\mathbf{j}^n\chi^m\mathbf{j}^{\dot{n}}\chi^{\dot{m}} &= \mathbf{j}^n\mathbf{j}^{\dot{n}}\mathbf{j}^{-\dot{n}}\chi^m\mathbf{j}^{\dot{n}}\chi^{\dot{m}} \\ &= \mathbf{j}^{n+\dot{n}}\chi^{(-1)^{\dot{n}}m+\dot{m}} = \mathbf{j}^{n'}\chi^{m'}\end{aligned}$$

for some integers n' and m' (here we made repeated use of the fact that $\mathbf{j}^{-1}\chi^m\mathbf{j} = \chi^{-m}$). Therefore $\mathsf{q}_1\mathsf{q}_2 \in B\mathcal{D}_d$ whenever q_1 and q_2 are. Hence $B\mathcal{D}_d$ is closed under quaternion multiplication, and by Theorem 5.53 it is a group.

Remark 5.111. The groups $B\mathcal{D}_n$ and $\mathcal{D}_{2n}$ are very close to eachother (with ρ playing the role of χ and $\mathbf{r}$ playing the role of $\mathbf{j}$)! The only difference is that in $\mathcal{D}_{2n}$ we have that $\mathbf{r}^2 = \mathbf{1}$ but in $B\mathcal{D}_n$ we have that $\mathbf{j}^2 = -1$. [II]

Example /Theorem 5.112. Every element of the form $\chi^a\mathbf{j} \in B\mathcal{D}_n$ has order 4.

There are several ways of proving this. I give three quick ones to help your familiarity with quaternions and their interaction with group theory. In the first part we use the following:

The first proof proves that in the bigger group $B\mathcal{D}_{2n}$ the set of elements $\chi^a\mathbf{j}$ (as a runs through the integers) are all conjugate to each other. Since conjugate elements have the same order (even if they are conjugate only in a larger group) this proves the result.

Proof. (**Proof 1**) Let $\alpha \in \mathbb{S}^3$ denote χ_{4n}. Note that $(\chi_{4n})^2 = \chi_{2n} = \chi$ and the pure part of α, α_P, is $\sin \frac{n\pi}{2}\mathbf{i}$ which is perpendicular to $\mathbf{j}$. Then we have that

$$\begin{aligned}\alpha\mathbf{j}\alpha^{-1} &= \alpha\mathbf{j}\alpha^{-1}\mathbf{j}^{-1}\mathbf{j} \\ &= \alpha\alpha^{+1}\mathbf{j} \qquad (\text{Since } \chi_P \perp \mathbf{j}) \\ &= \alpha^2\mathbf{j} \\ &= \chi\mathbf{j}.\end{aligned}$$

Similarly, for any integer a, $\alpha^a\mathbf{j}\alpha^{-a} = \chi^a\mathbf{j}$. Thus, in $\mathbb{S}^3$ all of the elements of the form $\alpha^a\mathbf{j}$ are conjugate to each other. As such, they all have the same order as $|\mathbf{j}|$, which is 4. ■

The second proof shows directly that $(\chi^a\mathbf{j})^2 = -1$.

Proof. (**Proof 2**) Again, we use the fact that $\mathbf{j}\chi^a\mathbf{j}^{-1} = \chi^{-a}$. Then we see

$$\begin{aligned}(\chi^a\mathbf{j})^2 &= (\chi^a\mathbf{j})(\chi^a\mathbf{j}) \\ &= (\chi^a\mathbf{j})(\chi^a\mathbf{j}^{-1}\mathbf{j}\mathbf{j}) \quad (\text{red term inserted}) \\ &= \chi^a(\mathbf{j}\chi^a\mathbf{j}^{-1})\mathbf{j}^2 \\ &= -\chi^a(\mathbf{j}\chi^a\mathbf{j}^{-1}) \qquad (\mathbf{j}^2 = -1) \\ &= -\chi^a\chi^{-a} \\ &= -1.\end{aligned}$$

Therefore $(\chi^a\mathbf{j})^2 \neq 1$ but $(\chi^a\mathbf{j})^4 = 1$. Therefore, we have proven that $\chi^a\mathbf{j}$ does indeed have order 4. ■

This next proof shows that the product $\chi^a\mathbf{j}$ is always in $\mathbb{S}^2$, so the product, like all elements of $\mathbb{S}^2$, has order 4.

Proof. (**Proof 3**) We have

$$\chi^a = \cos\frac{a\pi}{n} + \sin\frac{a\pi}{n}\mathbf{i} \Rightarrow \chi^a\mathbf{j} = \cos\frac{a\pi}{n}\mathbf{j} + \sin\frac{a\pi}{n}\mathbf{k}$$

which is pure. Moreover since $\mathbb{S}^3$ is closed under multiplication, we have that this product is still in $\mathbb{S}^3$. Hence the product is in the intersection $\mathbb{S}^3 \cap \mathbb{R}^3$ which is $\mathbb{S}^2$ (pure, unit quaternions are by definition in $\mathbb{S}^2$). ■

Exercise 5.68. Make a multiplication table for $B\mathcal{D}_3$. (If you make use of, Corollary 5.16, this could take less time!)

SOLUTION

Exercise 5.69. Write $\chi^2\mathbf{j}^2\chi\mathbf{j}\chi \in B\mathcal{D}_3$ as one of the elements in (14).

SOLUTION

(5) The sets of matrices $\mathbf{GL}_n(R)$ and $\mathbf{SL}_n(R)$ form groups under matrix multiplication for R a commutative ring, i.e. $R = \mathbb{Z}_n$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$. They are finite (only) in the case the $R = \mathbb{Z}_n$. This is clear because the total number of 2 by 2 matrices and coefficients in $\mathbb{Z}_n$ is n^4, so the subsets $\mathbf{GL}_2(\mathbb{Z}_n)$ and $\mathbf{SL}_2(\mathbb{Z}_n)$ must also be finite (see the exercise below, Exercise 5.72). Later we will make this more exact by computing the exact orders. For instance, much later, in Example/Theorem 10.98 we will see that for positive prime numbers,

$$\left|\mathbf{SL}_2\left(\mathbb{Z}_p\right)\right| = (p+1)\,p\,(p-1) \quad \text{and} \quad \left|\mathbf{GL}_2\left(\mathbb{Z}_p\right)\right| = (p+1)\,p\,(p-1)^2.$$

(The second fact comes from some basic Linear algebra, but we will prove this without using Linear Algebra). To tell you the truth, I do not know a general formula for the order of these groups when p is not a prime. We will be able to work a case by case basis though.

Exercise 5.70. First verify that $\mathbf{SL}_2(\mathbb{Z}_2)$ has 6 elements. Then make a multiplication table for this group.
Hint! Playing Sudoku may help you in this process.

Solution

Exercise 5.71. Verify that $\mathbf{SL}_2(\mathbb{Z}_3)$ has 24 elements by writing them all down. Construct as much of the multiplication table as you can until you get bored.

Solution

Exercise 5.72. Prove that TFAE

(1) R is infinite;

(2) $\mathbf{SL}_2(R)$ is infinite;

(3) $\mathbf{SL}_n(R)$ is infinite for any integer $n > 1$.

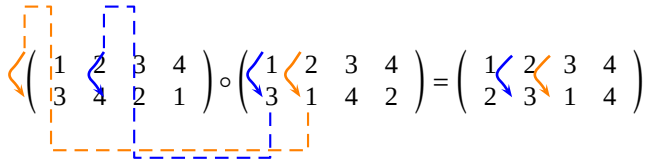
Solution

Hint! We elaborate the comments about this exercise given at the beginning of this section.

- If $\#R = m$, then there are a total of m^{n^2}, n by n matrices.
- $$\left\{\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} : r \in R\right\} \subseteq \mathbf{SL}_2(R).$$
- There is an identification of $\mathbf{SL}_n(R)$ with a subset of $\mathbf{SL}_{n+1}(R)$ for all positive integers n.

(6) In a sense, the most general kind of group is the set of bijections of any set X to itself (i.e. permutations of X) under composition of functions, $(\mathcal{S}_X, \circ)$. We've already discussed this group several times. The context of Theorem 2.57 is exactly that this is a group. (Of course we later generalized this theorem into The identity of this group is the identity function, $\mathbf{1}_X$ and the inverse of an element has the usual meaning as the inverse of the function. As we have already discussed in the section which started on page 36, for us, the most important example is the cases that the set is the standard set with n elements, $\overline{n} = \{1, \ldots, n\}$. In that same section we agreed to write $\mathcal{S}_n$ instead of $\mathcal{S}_{\overline{n}}$. In fact if you spend a few minutes playing with the set of bijections of the set, say, $\overline{3} = \{1, 2, 3\}$, and some other set with three elements — say $\{\spadesuit, \heartsuit, \clubsuit\}$ – you should be able to convince yourself that studying the bijections of the of both sets is the same. In fact if you were to write out the multiplication tables

for both groups they would be identical (up to the choice of the names). We also pointed out that the advantage of using numbers instead of anything else is because we will need to index any class of sets we deal with and the set $\overline{n}$ comes to us already indexed! Just go ahead and try to refer to a set with n elements without using the number n anywhere! Using the notation we have already discussed it is pretty straight forward to perform the operation (don't forget, composition from right to left). For instance if we follow the blue line we see that the first permutation (on the right!) takes 1 to 3, then the second one (on the left!) takes 3 to 2. Therefore (as indicated in the resulting permutation in blue) the composition takes 1 to 2. At the same time, if we follow the orange line we see that 2 goes to 1, then 1 goes to 3. Therefore, the net effect takes 2 to 3.[III]



Let's upgrade the table we started at the beginning of the chapter:

Group	$\mathbb{Z}_n$	$\mathcal{S}_n$	$\mathbf{SL}_2\left(\mathbb{Z}_p\right)$	$\mathcal{U}(\mathbb{Z}_n)$	$\mathbb{H}^\times$	$\mathbb{S}^3$	$\mathbb{C}^\times$	$\mathbb{S}^1$	$\mathbb{Z}$	$\mathcal{D}_n$	$B\mathcal{D}_n$
Order	n	$n!$	(p^3-p)	$\vartheta(n)$	∞	∞	∞	∞	∞	$2n$	$4n$
Abelian(?)	yes	no	no	yes	no	no	yes	yes	yes	no	no

Here are some things which do **not** form groups.

(1) The set of integers $\mathbb{Z}_n$ (including $\mathbb{Z}$) do not form a group under multiplication since 0 does not have a multiplicative inverse. In most cases there are more reasons than just this one why these algebraic systems do not form a group. For instance in $\mathbb{Z}$ the number 2 does not have a multiplicative inverse (the number $1/2$ is **not** an integer).

(2) The set of non-zero elements of $\mathbb{Z}_6$, denoted $\mathbb{Z}_6^\times$, does not form a group under multiplication since we have $[\,2\,]_6 \cdot [\,3\,]_6 = [\,0\,]_6$ and therefore multiplication isn't even a binary operation (because it's not closed). In fact $\mathbb{Z}_n^\times$ forms a group iff the number n is prime.

(3) The natural numbers, $\mathbb{N}$ under addition (or multiplication) are not a group since (say) 2 does not have an additive inverse (or a multiplicative inverse) inside of $\mathbb{N}$.

(4) The set $\{0, 1, \ldots, n-1\}$ (for any integer $n > 1$) does not form a group under addition (or multiplication) since $(n-1) + (n-1) > (n-1)$ and so addition isn't even a binary operation (to make this a group we need to take these dudes in $\mathbb{Z}_n$!)

(5) (**Calculus Example**) The set of invertible differentiable functions under the operation of composition does not form a group since x^3 is differentiable and invertible, but its inverse $\sqrt[3]{x}$ is not differentiable (its derivative at 0 does not exist). However, the set of differentiable functions with differentiable inverses does form a group under composition (this is known as the set of diffeomorphisms).

5.12 More

5.12.1 Fermat's Little Theorem

Here's an interesting application of the fact that for p a prime, $\mathcal{U}(\mathbb{Z}_p)$ is cyclic of order $p-1$:

[III] Don't forget, permutations are a special kind of function. And like all functions, composition goes (unfortunately) from right to left!

THEOREM 5.113 (Fermat's Little Theorem). *Let a be an integer and p a prime.*

(1) *If a is not divisible by p, then p divides $a^{p-1} - 1$;*

(2) *For any a, $a^p - a$ is divisible by p.*

PROOF.

(1) The first assumption guarantees that $[\, a \,]_p \in \mathcal{U}(\mathbb{Z}_p)$. The latter is a cyclic group of order $p - 1$. Therefore,

$$[\, a^{p-1} \,]_p = [\, 1 \,]_p \Rightarrow [\, a^{p-1} - 1 \,]_p = [\, 0 \,]_p \,.$$

Therefore $a^{p-1} - 1$ is divisible by p, which is what we wanted.

(2) We just showed that if a is not divisible by p, then $a^{p-1} - 1$ is. Therefore for any integer either a is divisible by p or $a^{p-1} - 1$ is. In either case the product

$$a \cdot (a^{p-1} - 1) = a^p - a$$

is divisible by p since at least one of the factors is divisible by p. ■

EXAMPLE 5.114. Since $10,432$ doesn't end in a 0 or a 5 it is not divisible by 5. Here the Corollary checks since

$$10432^4 - 1 = 11,843,234,091,237,375$$

is divisible by 5. For another example, consider $a = 142$. This number is not divisible by 7 (it has a remainder of 2) and

$$142^6 - 1 = 8,198,418,170,943 = 1,171,202,595,849 \cdot 7.$$

EXERCISE 5.73. (**Wilson's Theorem**) Prove that a number n is prime iff

$$(n-1)! \equiv_n -1.$$

SOLUTION

Hint!

- If n is a prime you will need to use the fact that $[\, -1 \,]_n$ is the unique element in $\mathcal{U}(\mathbb{Z}_n)$ which has order 2.
- If n is not prime, use the fact that there exists some $0 < m < n$ with $\gcd(n, m) > 1$.

6 Subgroups

Roughly speaking, a subgroup of a group is a group within a group. We have already seen this many times. For instance $\mathbb{Z}$ is a subgroup of $\mathbb{Q}$ which is a subgroup of $\mathbb{R}$ which is a subgroup of $\mathbb{C}$ which is a subgroup of $\mathbb{H}$ (all under addition). The same is true for the corresponding chain of invertible elements, $\{\pm 1\}$ is a subgroup of $\mathbb{Q}^\times$ which is a subgroup of $\mathbb{R}^\times$ which is a subgroup of $\mathbb{C}^\times$ which is a subgroup of $\mathbb{H}^\times$ (all under multiplication). Less trivial instances include the facts that $\mathbf{SL}_n(R)$ is a subgroup of $\mathbf{GL}_n(R)$ (under matrix multiplication) and $\mathbb{S}^1$ is a subgroup of $\mathbb{S}^3$ (under quaternion multiplication) which is a subgroup of $\mathbb{H}^\times$. Note that $\mathcal{U}(\mathbb{Z}_n)$, under any reasonable definition, shouldn't be a subgroup of $\mathbb{Z}_n$ since the operations for the two groups are very different.

We first introduce a more general notion of a subgroup inside of an algebraic system. After this short bit, for the rest of the chapter, we only study subgroups inside of a group.

6.1 Basic Definitions

Definition 6.1. Suppose that $(M, \star)$ be an algebraic system. Suppose that $H \subseteq M$ is a group under (the restriction of) $\star$ to H. Then we say that H is a *subgroup* of M.

For the most part, we will take M to be a group G. However, if M is not a group several pathologies can arise. For instance if we look at the algebraic system $(\mathbb{Z}_{10}, \cdot)$, we can see that (rather surprisingly) that the subset $\{[2]_{10}, [4]_{10}, [6]_{10}, [8]_{10}\}$ is a subgroup. Let's write out the table

$(\mathbb{Z}_{10}, \cdot)$	$[2]_{10}$	$[4]_{10}$	$[6]_{10}$	$[8]_{10}$
$[2]_{10}$	$[4]_{10}$	$[8]_{10}$	$[2]_{10}$	$[6]_{10}$
$[4]_{10}$	$[8]_{10}$	$[6]_{10}$	$[4]_{10}$	$[2]_{10}$
$[6]_{10}$	$[2]_{10}$	$[4]_{10}$	$[6]_{10}$	$[8]_{10}$
$[8]_{10}$	$[6]_{10}$	$[2]_{10}$	$[8]_{10}$	$[4]_{10}$

From this table we see that (highlighted in green) that (another surprise!) the element $[6]_{10}$ is the identity. This is because we can see that the top row is repeated below it in red (besides the $[6]_{10}$) and far left column is repeated to the right of it in red (under the $[6]_{10}$). With this strange identity, we see that $[2]_{10}^{-1} = [8]_{10}$, $[4]_{10}^{-1} = [4]_{10}$, and $[6]_{10}^{-1} = [6]_{10}$. Lastly, the operation is associative because it is the restriction of an associative operation. As such, for a subgroup H of M, we need to make a distinction between the identity of H, e_H, and the identity (that may or may not be there) in X, e_M. The same goes for inverse of an element, h_H^{-1} vs h_M^{-1}.

Another rather pathological example comes when we consider the algebraic system $(M, \star)$ from Example 4.11 back on page 216 which was given by the table

$\circ$	e	a	b
e	e	a	b
a	a	e	e
b	b	e	b

. Here we have each of

$$\{e\},\ \{e, a\},\ \{b\}$$

are subgroups of M. The first two listed have identity e, and the last listed has identity b.

6.2 Subgroups of Groups

To tell you the truth, the main reason why I introduced the notion of a subgroup of a general algebraic system was to make sure that you don't take the properties that we derive for subgroups of groups, which is the main interest of study for us, for granted. Unless specifically stated, whenever we say that H is a subgroup of an algebraic system G, this will mean that G is itself a group. In this situation we use the following notation

Notation 6.2. Suppose that H is a subgroup of a group G. Then we write $H \leqslant G$.

Let's begin by making a couple of obvious points for a subgroup H of G:

(1) If G is abelian, than so is H.

(2) If G is finite, then so is H.

(3) The set of all subgroups of a group forms a poset under $\leqslant$. This is because

Reflexive Any group is a subgroup of itself.

Anti-Symmetric If $H \leqslant K$ and $K \leqslant H$, then we have that $H \subseteq K$ and $K \subseteq H$ so $H = K$.

Transitive If $H \leqslant K$ and $K \leqslant G$, then H is a group and is a subset of G. Therefore $H \leqslant G$.

(4) For any group, G, $\{e_G\} \leqslant G$. (This is known as the trivial subgroup of G.)

So for any group G, both $\{e_G\}$ and the G itself are subgroups of G. We typically rule these cases out.

Definition 6.3. A subgroup of G other than the trivial subgroup, $\{e_G\}$ is called *non-trivial*. A subgroup of G besides G itself is called *proper*.

NOTATION 6.4. PNT is an acronym for proper, nontrivial.

Therefore, a PNT subgroup of G is a subgroup of G which is neither G itself or $\{e_G\}$.
The following lemma says that none of the pathologies which can arise in general actually arise in this situation.

LEMMA 6.5. *Suppose that $H \leqslant G$. Then*

(1) $e_H = e_G$.

(2) $h_H^{-1} = h_G^{-1}$

Therefore, we have the following:

PROPOSITION 6.6. *We have that $H \leqslant G$ iff*

(1) *If $h_1, h_2 \in H$, then $h_1 \cdot h_2 \in H$;* (**Closure under operation**)

(2) $e_G \in H$;

(3) *If $h \in H$, then h^{-1} is also in H.*(**Closure under inversion**) ■

We now improve upon this proposition.

6.3 SUBGROUP TESTS

DEFINITION 6.7. Let H be a non-empty subset of a group G.

(1) H passes the *two step subgroup test* if

TS1 H is closed under taking inverses. $(\forall h \in H, h^{-1} \in H)$
TS2 H is closed under multiplication. $(\forall h_1, h_2 \in H, h_1 h_2 \in H)$

(2) H passes the *one step subgroup test* if

OS1 H is closed under division. $(\forall h_1, h_2 \in H, h_1 h_2^{-1} \in H)$

(3) H passes the *one step subgroup test for finite groups* if

FOS1 H is closed under multiplication.

WARNING! 6.8. The names of these tests are a bit misleading since we must first verify that the set H is non-empty. In practice, most of the time, one does this by showing that the identity is an element of H.

THEOREM 6.9. *Suppose H is a non-empty subset of a group G. Then,* TFAE*:*

(1) *H is a subgroup of G;*

(2) *H passes the two-step subgroup test;*

(3) *H passes the one-step subgroup test.*

Moreover, if H happens to be finite, these are each equivalent to

(4) *H passes the one step subgroup test for finite groups.*

Before giving the proof we give a couple of examples.

EXAMPLE 6.10. This means that we can tell at a glance by looking at the multiplication table if a given subset of a group is a subgroup. It is even easier to tell when the original group is finite. For instance if we look at the set of even permutations on four letters, denoted as $\mathcal{A}_4$, consists of 12 elements, and we label them as follows:

$$1 = \begin{pmatrix}1&2&3&4\\1&2&3&4\end{pmatrix} \quad 2 = \begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix} \quad 3 = \begin{pmatrix}1&2&3&4\\3&4&1&2\end{pmatrix} \quad 4 = \begin{pmatrix}1&2&3&4\\4&3&2&1\end{pmatrix}$$
$$5 = \begin{pmatrix}1&2&3&4\\2&3&1&4\end{pmatrix} \quad 6 = \begin{pmatrix}1&2&3&4\\3&2&4&1\end{pmatrix} \quad 7 = \begin{pmatrix}1&2&3&4\\1&4&2&3\end{pmatrix} \quad 8 = \begin{pmatrix}1&2&3&4\\4&1&3&2\end{pmatrix}$$
$$9 = \begin{pmatrix}1&2&3&4\\3&1&2&4\end{pmatrix} \quad 10 = \begin{pmatrix}1&2&3&4\\1&3&4&2\end{pmatrix} \quad 11 = \begin{pmatrix}1&2&3&4\\2&4&3&1\end{pmatrix} \quad 12 = \begin{pmatrix}1&2&3&4\\4&2&1&3\end{pmatrix}.$$

(The only reason why I'm making this labeling is because I couldn't fit the multiplication table on the screen otherwise!) and the multiplication table for $\mathcal{A}_4$ looks as follows:

$\mathcal{A}_4$	1	2	3	4	5	6	7	8	9	10	11	12
1	1	2	3	4	5	6	7	8	9	10	11	12
2	2	1	4	3	7	8	5	6	12	11	10	9
3	3	4	1	2	8	7	6	5	10	9	12	11
4	4	3	2	1	6	5	8	7	11	12	9	10
5	5	6	7	8	9	10	11	12	1	2	3	4
6	6	5	8	7	11	12	9	10	4	3	2	1
7	7	8	5	6	12	11	10	9	2	1	4	3
8	8	7	6	5	10	9	12	11	3	4	1	2
9	9	10	11	12	1	2	3	4	5	6	7	8
10	10	9	12	11	3	4	1	2	8	7	6	5
11	11	12	9	10	4	3	2	1	6	5	8	7
12	12	11	10	9	2	1	4	3	7	8	5	6

Since every entry corresponds to an element of $\mathcal{A}_4$, the subset $\mathcal{A}_4$ is closed under the operation and so it satisfies **FOS**. Since it is also finite (it only has 12 elements!) it passes the one step subgroup test for finite groups. Therefore $\mathcal{A}_4$ is a subgroup of $\mathcal{S}_4$ (i.e. $\mathcal{A}_4 \leqslant \mathcal{S}_4$). In particular, $\mathcal{A}_4$ is a group. (We will see later that $\mathcal{A}_n$ is always a group!) If we were gluttons for punishment, and we enjoyed checking things that we are guaranteed to get for free from the general theory, we could also check that the inverse of every element in $\mathcal{A}_4$ is also in $\mathcal{A}_4$ (now we are applying the two step subgroup test, which we don't need to do since we are looking at finite crap here!) For instance looking across the 5 row we see a 1 in the 9 column. This means that $5^{-1} = 9$. Or if you prefer you could also look down the 5 column or across the 9 row or even down the 9 column to make the same conclusion. However, part of theorem says that there is no need to do this: the second a finite subset is closed under multiplication it is a group, and hence every element is guaranteed to have an inverse. [1] At the same time we can further pull out the subset of $\mathcal{A}_4$ given by $\{1, 2, 3, 4\}$ and observe that this set is also closed under multiplication and, as such, is also a subgroup of $\mathcal{S}_4$ (or if you prefer $\mathcal{A}_4$). Other subgroups include

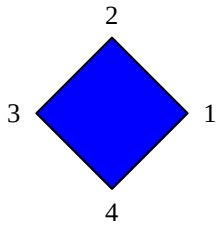
$$\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5, 9\}, \{1, 6, 12\}, \{1, 7, 10\}, \{1, 8, 11\}.$$

It is rather surprising that these are actually the only PNT subgroups of $\mathcal{A}_4$ (i.e. the only other subgroups of $\mathcal{A}_4$ are either all of $\mathcal{A}_4$ or only the identity). You will prove this in the chapter on Lagrange's theorem. There you will be aided by the fact that we will prove the amazing fact that the order of any subgroup must divide the order of the group. There is however several more subgroups of $\mathcal{S}_4$. For instance, we could check directly that the set of permutations

$$\begin{pmatrix}1&2&3&4\\1&2&3&4\end{pmatrix} \quad \begin{pmatrix}1&2&3&4\\2&3&4&1\end{pmatrix} \quad \begin{pmatrix}1&2&3&4\\3&4&1&2\end{pmatrix} \quad \begin{pmatrix}1&2&3&4\\4&1&2&3\end{pmatrix}$$
$$\begin{pmatrix}1&2&3&4\\1&4&3&2\end{pmatrix} \quad \begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix} \quad \begin{pmatrix}1&2&3&4\\3&2&1&4\end{pmatrix} \quad \begin{pmatrix}1&2&3&4\\4&3&2&1\end{pmatrix}$$

form a subgroup of $\mathcal{S}_4$ which we denote as $\widetilde{\mathcal{D}_{4,1}}$. Indeed, these permutations are exactly the set of permutations that one gets when $\mathcal{D}_4$, the symmetries of a square, permutes the 4 vertices of the square labeled as follows:

[1] You can ignore the coloring for the time being, but we will explain in more detail what the coloring refers to. For now you can simply observe that anything colored blue times anything colored green is colored red, etc.



EXAMPLE 6.11. In Example 5.5 we proved directly (by verify all 4 group axioms) that $\mathbb{S}^3$ is a group (hence a subgroup of $\mathbb{H}^\times$). However, simply using 3.101**(1)** which says that $\mathbb{S}^3$ is closed under both multiplication and inversion, we have proven (much more quickly) that $\mathbb{S}^3 \leqslant \mathbb{H}^\times$, and in particular $\mathbb{S}^3$ is a group.

Clearly, any subgroup of G must pass any of the subgroup tests. So this theorem is really telling us that a subset of a group which passes either the one step subgroup test or the two step subgroup test, or is a finite set passing the one step for finite is a group. Often this is the most straight-foward way of actually showing that something is a group in the first place, i.e. the fastest way to show that a given set is to show that it's a subgroup of something which is obviously a group.

WARNING! 6.12. The names of these tests are a bit misleading since in practice, one often needs to perform an extra step to show that the subset is non-empty in the first place. Ironically, in theoretical applications one does this by showing that the identity is in the stated subset in the first place (which kinda sucks cuz that's exactly what is saved by doing the two step subgroup test in the first place!) Thus you might find it helpful to insert a step 0 into each of these tests which states and proves that "*H is non-empty*".

PROOF. (**Proof of Theorem 6.9**). We already have that $H \leqslant G$ implies that H passes all three tests: **TS**, **OS**, and **FOS**.

TS⇒⩽ Suppose $H \leqslant G$ passes the two step subgroup test. By Proposition 6.6 we only need to prove that $e_G \in H$. But since H is non-empty there is some element $h \in H$. By **TS2**, which holds since we are assuming that H passes the two step subgroup test, we know that $h^{-1} \in H$. Thus by **TS1** we have that $hh^{-1} \in H$. But $hh^{-1} = e_G$, and so $e_G \in H$ as we wanted.

OS ⇒ TS Suppose that H is a subset of a group G which passes the one step subgroup test. We show that H also passes the two step subgroup test, and is hence a subgroup of G.

TS1 Since $H \neq \varnothing$ we have at least one element $x \in H$. And by **OS1** we have that $xx^{-1} \in H$. But this product equals e_G so we have shown that $e_G \in H$. Now by **OS1** we know that for all $h \in H$, $e_G h^{-1} \in H$. But this product is simply h^{-1}. Thus H is closed under inverses.

TS2 Let $h_1, h_2 \in H$. We must show that $h_1 h_2 \in H$. Since we just proved **TS1** we have that $h_2^{-1} \in H$ as well. Hence by **OS1** we have that $h_1 \left(h_2^{-1}\right)^{-1} \in H$. But this product equals $h_1 h_2$. Thus H is closed under multiplication, i.e. **TS2** has been proven.

FOS⇒⩽ This is precisely the content of Theorem 5.53 back on page 277. ∎

THEOREM 6.13. *Let $g \in G$. Then the cyclic group generated by g, $\langle g \rangle$, is*

(1) *a group*

(2) *abelian*

(3) *of order $|g|$.*

REMARK 6.14. This is a loaded theorem, it really says three different things:

(1) $\langle g \rangle$ doesn't have a crappy name because it is a group.

(2) Any two elements of $\langle g \rangle$ commute.

(3) The set $\langle g \rangle$ has $|g|$ elements in it.

PROOF.

(1) Since $\langle g \rangle$ is non-empty (it contains $g^1 = g$), we can (and will!) use the one step subgroup test.

OS1 Let $h, k \in \langle g \rangle$. Then there are integers n, m so that $h = g^n$ and $k = g^m$. Then

$$k^{-1} = (g^n)^{-1} = g^{-n} \quad \text{and} \quad hk^{-1} = g^m g^{-n} = g^{m-n}$$

which is a power of g. Hence $hk^{-1} \in \langle g \rangle$ as desired.

(2) We gave the argument that $\langle g \rangle$ is abelian immediately after giving the definition. We lay the argument out here a bit more carefully. Let $x, y \in \langle g \rangle$. By definition, there are integers m and n so that $x = g^m \in \langle g \rangle$ and $y = g^n \in \langle g \rangle$. As such, we have that

$$xy = g^m g^n = g^{m+n} = g^{n+m} = g^n g^m = yx.$$

Therefore, any pair of elements of $\langle g \rangle$ commute.

(3) The statement about the order is exactly the statement in Theorem 5.61. ■

In a sense $\langle g \rangle$ is the smallest group which contains g. Indeed, by closure, if a subgroup $H \leqslant G$ contains the element g, then it must contain every power of g. But since $\langle g \rangle$ consists of the powers of g, it must contain all of $\langle g \rangle$.

PROPOSITION 6.15. *If $H, K \leqslant G$, then $H \cap K \leqslant G$.*

PROOF. For starters, this intersection is non-empty (why?). So we can use a subgroup test. For variety, we use the 2 step version. Let $a, b \in H \cap K$.

TS1 Since $a, b \in H \cap K$ we have $a, b \in H$ and $a, b \in K$. Therefore by closure of H we have that $ab \in H$ and by closure of K we have that $ab \in K$. Therefore $ab \in H \cap K$.

TS2 Since $a \in H$ we have that $a^{-1} \in H$ (by closure of inverses in H) and since $a \in K$ we also have that $a^{-1} \in K$. Therefore, a^{-1} is in both H and K and, as such, is in the intersection, $H \cap K$. ■

WARNING! 6.16. Try to prove that in this situation $H \cup K$ is a subgroup. Most students screw this up! In fact: **IT IS NOT TRUE THAT THE UNION OF TWO SUBGROUPS IS A SUBGROUP**. Moreover it is almost **NEVER** true. For instance, $\{\mathbf{1}, \mathbf{r}\}$ and $\{\mathbf{1}, \rho^2\}$ are subgroups of $\mathcal{D}_4$ but their union is $\{\mathbf{1}, \mathbf{r}, \rho^2\}$ which is not a subgroup since $a^2 b$ is not in this union implying that this union is not closed under the operation. If you did indeed prove something that was false go back and try to identify what your mistake was! (Most students make the same mistake at the same spot!)

Compared to Proposition 6.15, it is only slightly harder to prove the following generalization:

THEOREM 6.17. *Suppose that $\{H\}_{H\in\mathcal{A}}$ is a non-empty collection of subgroups of a group G. Then*

$$\bigcap_{H\in\mathcal{A}} H$$

is also a subgroup of G. (In words the intersection of any non-empty collection of subgroups of a group is also a subgroup.)

PROOF. Exercise 6.1. ∎

EXERCISE 6.1. Prove Theorem 6.17.

SOLUTION

Let $S \subseteq G$ be any old subset of a group (not necessarily a subgroup). Then there exists at least one subgroup of G which contains G, namely the entire group G itself. As such if we take the intersection of all subgroups of G which contains S, this intersection is over a non-trivial collection of sets each of which contains S. Therefore, by the theorem, Theorem 6.17 just above,

$$\bigcap_{S\subseteq H\leqslant G} H \tag{1}$$

is a subgroup of G which contains S. Therefore we can make the following definition:

DEFINITION 6.18. For any subset $S \subseteq G$ we define the group *generated* by S as the intersection in Equation (1). We denote this as $\langle S \rangle$.

While this definition is easy to state it is essentially useless in practice. We will spend an entire chapter, Chapter 9, on this concept where we will see better version of this concept For now we can think of $\langle S \rangle$ as the smallest subgroup of G which contains S. We only introduce the concept here because it allows us to introduce more examples. For now we content ourselves with the following, which relates this new concept with what we previously defined as $\langle s \rangle = \{s^n : n \in \mathbb{Z}\}$. It should be fairly intuitive since $\langle s \rangle$ certainly is the smallest subgroup of G which contains s.

THEOREM 6.19. *Let $s \in G$. Then $\langle s \rangle = \langle \{s\} \rangle$.*

PROOF. Suppose that $H \leqslant G$ contain s. Then, by closure, H also contains every power of s. As such, by definition, H contains $\langle s \rangle$. Therefore $\langle s \rangle \subseteq \langle \{s\} \rangle$. But we have seen that $\langle s \rangle$ is a subgroup of G. As such, to get $\langle \{s\} \rangle$ we are intersecting a collection of subgroups of G, one of which is $\langle s \rangle$. Hence the opposite inclusion also holds. ∎

As such it makes sense to not right brackets whenever the set is explicitly written. For instance, we can write $\langle a, b \rangle$ instead of the more cumbersome $\langle \{a, b\} \rangle$.

6.4 Examples

Here are some basic examples of subgroups.

Example 6.20.

(1) Corollary 3.101 shows that $\mathbb{S}^1$ passes the two step subgroup test and is hence a subgroup.[I]

(2) Exercise 3.19 shows that $\{\{n\}\} \preccurlyeq \mathbb{S}^1$. (It has order n.)

Example 6.21. We have already discussed subgroups of the integers, $\mathbb{Z}$. Indeed, in Example 3.22 back on page 80 we showed that for every integer $n \in \mathbb{Z}$, the set

$$n\mathbb{Z} = \{nz : z \in \mathbb{Z}\}$$

of multiples of n forms a subgroup. Note that these are the cyclic subgroups of $\mathbb{Z}$. In fact, later in Lemma 3.35 we actually showed that these were the only subgroups of $\mathbb{Z}$. At the same time, for any cyclic group $\langle g \rangle$ we can easily see that for any subgroup $H \preccurlyeq \langle g \rangle$ that

$$\{m \in \mathbb{Z} : g^m \in H\} \tag{2}$$

is a subgroup of $\mathbb{Z}$ (since this set is equal to $\mathbf{E}_g^{-1}(H)$ this is a special case of Theorem 8.12**(2)** coming up on page 377). Therefore as a subgroup of $\mathbb{Z}$, the set in (2) is equal to $d\mathbb{Z}$ for some integer d. This means that

$$H = \left\{\left(g^d\right)^m : m \in \mathbb{Z}\right\}$$

and so $H = \left\langle g^d \right\rangle$. This proves that every subgroup of a cyclic group is cyclic. We record these facts as a theorem:

Theorem 6.22.

(1) *Every subgroup of $\mathbb{Z}$ is cyclic and equals $d\mathbb{Z}$ for some integer $d \in \mathbb{Z}$.*

(2) *Every subgroup of a cyclic group is cyclic.* ∎

Corollary 6.23. *If C is a cyclic group of order n, then for every $d \mid n$ there exists a unique subgroup of C of order d and it happens to be*

$$C[d].$$

Moreover no other subgroups of C exist.

Proof. Suppose $g \in C$ is a generator. Then, we just proved that every subgroup of C has the form $\left\langle g^{d'} \right\rangle$ for some divisor of n. If this group is to have order d, then $\frac{n}{d'}$ needs to equal d. Moreover, by Exercise 5.58 back on page 292,

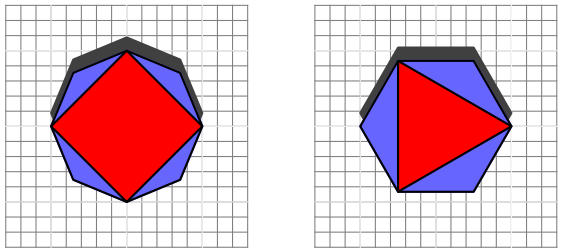
$$\left\langle g^{d'} \right\rangle = C[d].$$ ∎

Example 6.24. We have $\mathcal{D}_n \preccurlyeq \mathcal{D}_{qn}$ for any (positive) integers n and q. This is because the rotation

$$\varrho_{2\pi/qn}^q = \varrho_{2\pi/n}.$$

For instance we have that $\mathcal{D}_4 \preccurlyeq \mathcal{D}_8$ and $\mathcal{D}_3 \preccurlyeq \mathcal{D}_6$. Indeed this is easy to see since the relevant n-gons can be inscribed in such a way that symmetries induce symmetries. See Figure 6.24

[I] It is actually a subgroup of each of $\mathbb{C}^\times$, $\mathbb{H}^\times$ and $\mathbb{S}^3$.



EXAMPLE 6.25. The set of invertible matrices of the form

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with $a + b = 1$ and $c + d = 1$ is a subgroup of $\mathbf{GL}_2(\mathbb{Z})$. We denote the set by H_1. Clearly H_1 is non-empty since it contains the identity matrix, $\mathbf{I}_2$. Therefore, we can prove that it is a subgroup by applying the two step subgroup test. Let's consider two arbitrary matrices

$$\mathbf{A} := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad \mathbf{X} := \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

in H_1. (All this means is that we have $a + b = c + d = x + y = z + w = 1$.)

TS1 We need to show that $\mathbf{A} \cdot \mathbf{X} \in H_1$. We have that

$$\mathbf{A} \cdot \mathbf{X} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} ax + bz & ay + bw \\ cx + dz & cy + dw \end{pmatrix}.$$

Now we need to see that the rows of this product still sum to 1.

$$(ax + bz) + (ay + bw) = a(x + y) + b(z + w) = a \cdot 1 + b \cdot 1 = a + b = 1$$

and similarly

$$(cx + dz) + (cy + dw) = c(x + y) + d(z + w) = c \cdot 1 + d \cdot 1 = c + d = 1.$$

This proves that $\mathbf{A} \cdot \mathbf{X} \in H_1$, and H_1 is therefore closed under multiplication.

TS2 Now we must prove that $\mathbf{A}^{-1} \in H_1$. We use a bit of trickery to do this. We start by rewriting the determinant of $\mathbf{A}$ using the given relations:

$$\det \mathbf{A} = ad - bc = a(1 - c) - (1 - a)c = a - ac - c + ac = a - c.$$

This means the inverse is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{a - c} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

As such, we still need to prove that both

$$\frac{d - b}{a - c} \quad \text{and} \quad \frac{a - c}{a - c}$$

are equal to 1. The second one is silly. As to the first, note that

$$d - b = d(a + b) - b(c + d) = ad + bd - bc - bd = ad - bc = a - c.$$

Therefore, the ratio is 1 and H_1 is closed under inversion.

EXAMPLE /THEOREM 6.26. For p a prime integer, we claim that $\mathbb{Z}_p$ contains no PNT subgroups (i.e. that the only subgroup of $\mathbb{Z}_p$ are the trivial subgroup containing only the identity, and the entire group itself). It will actually turn out that these are the only groups where this is true. Indeed suppose that

$$H \leqslant \mathbb{Z}_p.$$

We can assume that H is non-trivial and hence it contains an element $[\,a\,]_p \in H$ which does not equal the identity. Moreover by closure of H under the operation from $\mathbb{Z}_p$ (addition) we know that $\left\langle [\,a\,]_p \right\rangle \leqslant H$. We show that $\left\langle [\,a\,]_p \right\rangle = \mathbb{Z}_p$ which implies that H also equals $\mathbb{Z}_p$. Indeed let $[\,m\,]_p \in \mathbb{Z}_p$. Then since p is prime $[\,a\,]_p$ has a multiplicative inverse $[\,n\,]_p \in \mathbb{Z}_p$ for some integer $n \in \mathbb{Z}$. This means that $n\,[\,a\,]_p = [\,n\,]_p\,[\,a\,]_p = [\,1\,]_p$. But $n\,[\,a\,]_p$ is $[\,a\,]_p$ added to itself n times. Therefore, by closure of $\left\langle [\,a\,]_p \right\rangle$ under addition, we have that $[\,1\,]_p \in \left\langle [\,a\,]_p \right\rangle$. Lastly, since

$$[\,m\,]_p = m\,[\,1\,]_p = mn\,[\,a\,]_p$$

we can conclude that $[\,m\,]_p \in \left\langle [\,a\,]_p \right\rangle$ as desired.

EXAMPLE 6.27.

(1) I claim that the set of non-zero complex matrices of the form

$$\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}$$

forms a subgroup of $\mathbf{GL}_2(\mathbb{C})$.[II] Let's denote the set in question as $\mathcal{H}$. For starters, let's say in words what it means for a matrix to be an element of $\mathcal{H}$: the conjugate upper left hand entry of a matrix in $\mathcal{H}$ must equal the lower right hand entry and negative the conjugate of the upper right hand entry must equal the lower left hand entry.

We apply the two step subgroup test to $\mathcal{H}$.

- Let

$$\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} z_1 & w_1 \\ -\overline{w}_1 & \overline{z}_1 \end{pmatrix}$$

be arbitrary matrices in $\mathcal{H}$. Then

$$\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} \cdot \begin{pmatrix} z_1 & w_1 \\ -\overline{w}_1 & \overline{z}_1 \end{pmatrix} = \begin{pmatrix} zz_1 - w\overline{w}_1 & zw_1 + w\overline{z}_1 \\ -\overline{w}z_1 - \overline{z}w_1 & -\overline{w}w_1 + \overline{z}z_1 \end{pmatrix}$$

is in $\mathcal{H}$ since the conjugate of the upper left entry of this product is

$$\overline{zz_1 - w\overline{w}_1} = \overline{z}\overline{z}_1 - \overline{w}w_1$$

which is the lower right entry. Meanwhile, negative the conjugate of the upper right entry is

$$-\overline{zw_1 + w\overline{z}_1} = -\overline{z}\overline{w}_1 - \overline{w}z_1$$

which is indeed the lower left entry. Hence whenever two matrices are in $\mathcal{H}$ so is there product and $\mathcal{H}$ is closed under multiplication.

[II] Dudes like

$$\begin{pmatrix} 1+i & 2+3i \\ -2+3i & 1-i \end{pmatrix}$$

are in this set.

- Let

$$\mathbf{A} = \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}$$

be an arbitrary matrix in $\mathcal{H}$. The assumption is that every matrix in $\mathcal{H}$ is non-zero. This means that either z or w must be non-zero. But since the determinant of $\mathbf{A}$ is

$$\|z\|^2 + \|w\|^2$$

we must have that the determinant is non-zero and the inverse exists. However we still need to check that it's an element of $\mathcal{H}$. Now, the inverse of the matrix $\mathbf{A}$ is

$$\frac{1}{\|z\|^2 + \|w\|^2} \cdot \begin{pmatrix} \overline{z} & -w \\ \overline{w} & z \end{pmatrix} = \begin{pmatrix} \frac{\overline{z}}{\|z\|^2+\|w\|^2} & \frac{-w}{\|z\|^2+\|w\|^2} \\ \frac{\overline{w}}{\|z\|^2+\|w\|^2} & \frac{z}{\|z\|^2+\|w\|^2} \end{pmatrix}.$$

Now bearing in mind that $\|z\|^2 + \|w\|^2$ is real and hence conjugating it does squat, the conjugate of the upper left entry of this matrix is

$$\overline{\frac{\overline{z}}{\|z\|^2 + \|w\|^2}} = \frac{z}{\|z\|^2 + \|w\|^2}$$

which is the lower right entry while the negative conjugate of the upper right hand entry is

$$-\overline{\frac{-w}{\|z\|^2 + \|w\|^2}} = \frac{\overline{w}}{\|z\|^2 + \|w\|^2}$$

which is the lower left entry.

Thus $\mathcal{H}$ passes the two step subgroup test and is therefore a subgroup (and hence a group).

(2) The subset of $\mathcal{H}$ consisting of matrices of determinant 1, denoted as $S\mathcal{H}$ is also a subgroup of $\mathbf{GL}_2(\mathbb{C})$ since we have the equality

$$S\mathcal{H} = \mathcal{H} \cap \mathbf{SL}_2(\mathbb{C})$$

which shows that it's the intersection of two subgroups of $\mathbf{GL}_2(\mathbb{C})$ and Theorem 6.15 applies.

Exercise 6.2. Suppose $H' \leqslant H$ and $K' \leqslant K$. Prove that $H' \times K' \leqslant H \times K$.

Solution

Exercise 6.3. Prove that for any abelian group A and any integer n prove that $A[n] \leqslant A$.

Solution

Exercise 6.4. Prove that $\mathcal{D}_4[2]$ is not a subgroup of $\mathcal{D}_4$.

Solution

The following exercise introduces a favorite group of many physicists. In this exercise, $\mathbf{U}_n(\mathbb{C})$ has nothing to do with the set of units from Definition 4.29 back on page 225.

Exercise 6.5. Let $\mathbf{A}$ be an n by n matrix with complex coefficients (i.e. an element of $\mathbf{M}_n(\mathbb{C})$). The conjugate of $\mathbf{A}$, $\overline{\mathbf{A}}$, is the matrix one gets if they were to replace all entries of $\mathbf{A}$ with their complex conjugate. The *hermitian conjugate* of $\mathbf{A}$, denoted $\mathbf{A}^\dagger$, is the conjugate of the transpose of $\mathbf{A}$:

$$\mathbf{A}^\dagger \stackrel{\text{def}}{=} \overline{(\mathbf{A}^\top)}.$$

$\mathbf{A}$ is called *unitary* if its inverse is its hermitian conjugate:

$$\mathbf{A} \cdot \mathbf{A}^\dagger = \mathbf{I}_n.$$

The set of all (respectively of determinant 1) n by n unitary matrices is denoted $\mathbf{U}_n(\mathbb{C})$ (resp. $\mathbf{SU}_n(\mathbb{C})$) and is called the *unitary group* (resp. the *special unitary group*) of degree n.

(1) Show that $\mathbf{SU}_n(\mathbb{C}) \preccurlyeq \mathbf{U}_n(\mathbb{C}) \preccurlyeq \mathbf{GL}_n(\mathbb{C})$.

(2) For any element of $\mathbf{U}_n(\mathbb{C})$ prove that $\det \mathbf{A} \in \mathbb{S}^1$.

(3) Show that $\mathbf{SU}_2(\mathbb{C}) = S\mathcal{H}$ where $S\mathcal{H}$ is the group just defined in Example 6.27.
Hint! There are two obvious ways to start this. One works right away, the other, to my knowledge, takes a very long time. If this doesn't "fall out" right away perhaps try another approach!

Hint! See the proofs about $\mathbf{O}_n(\mathbb{R})$ in C.2 on page 757 to get an idea how to proceed.

SOLUTION

Exercise 6.6. Show that for any prime number p

$$\left\{ \begin{pmatrix} [a]_{p^2} & [b]_{p^2} \\ [0]_{p^2} & [1]_{p^2} \end{pmatrix} : [a]_p = [1]_p \right\}$$

forms a group of order p^3.

Clarification! Of course this is a two part problem. Not only do you need to show that this dude is a group, you need to show that it has p^3 elements. [III]

SOLUTION

Notation 6.28. We denote the group from the last example as which we denote as Af_p.

[III] For example,

$$\begin{pmatrix} [4]_9 & [2]_9 \\ [0]_9 & [1]_9 \end{pmatrix} \quad \begin{pmatrix} [7]_9 & [0]_9 \\ [0]_9 & [1]_9 \end{pmatrix} \quad \begin{pmatrix} [7]_9 & [6]_9 \\ [0]_9 & [1]_9 \end{pmatrix}$$

are several elements of Af_3.

Let R be a commutative ring. Let $\mathbf{A}$ be an n by n matrix with coefficients in R (i.e. an element of $\mathbf{M}_n(R)$). A vector $\mathbf{v} \in R^n$ is called an *eigenvector* of $\mathbf{A}$ if there exists a $\lambda \in R$ so that

$$\mathbf{A} \cdot \mathbf{v} = \lambda \mathbf{v}.^{IV}$$

Exercise 6.7.

Solution

(1) Now fix a $\mathbf{v} \in R^n$. Prove that the set of invertible matrices (resp. matrices with determinant 1) which have $\mathbf{v}$ as an eigenvector form a subgroup of $\mathbf{GL}_n(R)$ (resp. $\mathbf{SL}_n(R)$) which we denote as $\mathbf{E}_\mathbf{v}$ (respectively $\mathbf{SE}_\mathbf{v}$).

(2) Prove that $\mathbf{E}_{(1,1)}$ (where $(1,1)$ is considered an element of $\mathbb{Z}^2$) contains the subgroup from 6.25. Can you describe $\mathbf{E}_{(1,1)}$ in that context?

(3) Describe the groups $\mathbf{SE}_{(1,0)}$ where

(a) $(1,0)$ is considered as an element of $\mathbb{Z}^2$,

(b) $(1,0)$ is considered as an element of $\mathbb{R}^2$.

(c) $(1,0)$ is considered as an element of $\mathbb{Z}_3^2$ (i.e. $([1]_3, [0]_3)$). In this part explicitly write down all elements
Hint! There are 6.

Example 6.29. Let X be a set and let $x \in X$. Then the set

$$(\mathcal{S}_X)_x := \{\sigma \in \mathcal{S}_X : \sigma(x) = x\}$$

is called the *stabilizer* of x. It is a subgroup of $\mathcal{S}_X$ since for any $\sigma, \tau \in (\mathcal{S}_X)_x$ we have

- We have $\sigma \circ \tau$ is also in $(\mathcal{S}_X)_x$ since
$$(\sigma \circ \tau)(x) = \sigma(\tau(x)) = \sigma(x) = x.$$
- We also have that $\sigma^{-1} \in (\mathcal{S}_X)_x$ since we have that
$$x = \mathbf{1}_X(x) = \left(\sigma^{-1} \circ \sigma\right)(x) = \sigma^{-1}(\sigma(x)) = \sigma^{-1}(x).$$

Exercise 6.8. Explain why it's reasonable to identify $(\mathcal{S}_X)_x$ with $\mathcal{S}_{X \setminus \{x\}}$.

Solution

Exercise 6.9. Let's generalize the previous example a bit. Let $A \subseteq X$. Define

$$(\mathcal{S}_X)_A := \{\sigma \in \mathcal{S}_X : \sigma(A) = A\}.$$

(For a finite subset A this is more simply written as the set of $\sigma \in \mathcal{S}_X$ which satisfy for all $a \in A$, $\sigma(a) \in A$).

Solution

[IV] The scalar λ is known as an eigenvalue of T.

(1) Show that this is also a subgroup of $\mathcal{S}_X$ (called the stabilizer of A).

(2) Show that

$$(\mathcal{S}_X)_A = (\mathcal{S}_X)_{X\setminus A}\,.$$

(3) Explicitly write down the elements of $(\mathcal{S}_4)_{\{1,3\}}$
Hint! The permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$$

is in there since it takes both 1 and 3 to either 1 or 3, i.e. calling this permutation σ,

$$\sigma(\{1,3\}) = \{1,3\}.$$

6.5 The Center and the Centralizer

Its particularly nice to be able to derive an entire family of subgroups instead of only deriving subgroups in very specific cases. Group theory has many instances of this. We start with a few of the most basic.

Definition 6.30. Let G be a group. Suppose $H \leqslant G$ and $a \in G$. Define the set

$$C_H(a) \stackrel{\text{def}}{=} \{h \in H : ha = ah\}.$$

More generally, for $S \subseteq G$ (S a subset, not necessarily a subgroup) we define

$$C_H(S) \stackrel{\text{def}}{=} \{h \in H : \forall a \in S, ha = ah\}.$$

We will typically be concerned with the case that $H = G$ and S is either all of G or S is a single element. Here are a couple of special cases of importance:

SubDefinition 6.31.

(1) The *centralizer* of a in G is defined to be $C_G(a)$.

(2) The *center* of G, denoted $\mathcal{C}(G)$, is defined to be $C_G(G)$.

That is the centralizer of an element $g \in G$ is the set of elements which commute g, and the center is the set of guys who commute with *every* element of G. In symbols

$$\mathcal{C}(G) \stackrel{\text{def}}{=} C_G(G) = \{g \in G : \forall a \in G, ga = ag\}.$$

Exercise 6.10. Prove that

$$\mathcal{C}(G) = \bigcap_{s \in G} C_G(s)$$

and more generally that

$$C_H(S) = \bigcap_{s \in S} C_H(s).$$

(This problem is nearly trivial! Just say both sides in words!)

Solution

Exercise 6.11. Prove that $z \in \mathcal{C}(G)$ iff $C_G(z) = G$ (once again, this is simply a rephrasing of definitions).

Solution

EXERCISE 6.12.

(1) For $H \leqslant G$ and $S \subseteq G$ prove that $C_H(S) = C_G(S) \cap H$.

(2) Show, by way of example, that it is not true in general that $\mathcal{C}(H) = \mathcal{C}(G) \cap H$.
Hint! It's hard to pick an example where you go wrong! Just try a couple of examples!

SOLUTION

EXERCISE 6.13. Let $g \in G$ and $z \in \mathcal{C}(G)$. Prove that $C_G(g) = C_G(zg)$.

SOLUTION

EXERCISE 6.14. For two groups H and K prove that

$$\mathcal{C}(H \times K) = \mathcal{C}(H) \times \mathcal{C}(K).$$

SOLUTION

EXERCISE 6.15. Show that $g \in \mathcal{C}(G)$ iff $\#\mathrm{cc}_G(g) = 1$ (see Definition 5.101 back on page 299).

SOLUTION

EXAMPLE / THEOREM 6.32. For subsets $T, S \subseteq G$ we have $T \subseteq C_G(S)$ iff $S \subseteq C_G(T)$.[I]

PROOF. Then

$$\begin{aligned} T \subseteq C_G(S) &\Leftrightarrow \forall t \in T, t \in C_G(S) \\ &\Leftrightarrow \forall t \in T, \forall s \in S, t \text{ and } s \text{ commute.} \\ &\Leftrightarrow \forall s \in S, \forall t \in T, t \text{ and } s \text{ commute.} \\ &\Leftrightarrow \forall s \in S, s \in C_G(T) \\ &\Leftrightarrow S \subseteq C_G(T). \end{aligned}$$

■

For instance, a special case of this is $g \in \overset{C_G(G)}{\overset{\|}{\mathcal{C}}}(G)$ (i.e. $\{g\} \subseteq C_G(G)$) iff $G \subseteq C_G(g)$ i.e. $G = C_G(g)$, which was Exercise 6.11.

PROPOSITION 6.33. *For any subset $S \subseteq G$ and subgroup $H \leqslant G$, we have that $C_H(S) \leqslant G$.*

Since the center and the centralizer are special cases of this we have

[I] This really boils down to symmetric, reflexive relations, which $g \approx \overline{g}$ iff g and $\overline{g}$ commute certainly is an example of.

COROLLARY 6.34. *The centralizer of any element and the center of any group are both subgroups.* ∎

REMARK 6.35. In words, we are saying that if a and b commute with a fixed element g, then ab also commutes with g. The idea is that we can pass g by a and b one at a time.

PROOF. (**Proof of Proposition 6.33**) Since $C_H(S)$ is the intersection of centralizers of single points by Exercise 6.10, it's enough to prove that the centralizer of a single element of a group is a subgroup of that group. Let $s \in G$ be arbitrary. Since any element of a group commutes with itself, $C_G(s) \neq \emptyset$, and we can apply the two step subgroup test to it. Let g and $\overline{g} \in C_G(s)$ be arbitrary. (That is we are assuming every that g and $\overline{g}$ commute with s.)

TS1 By definition, to show that $g\overline{g} \in C_G(s)$ we must show that $g\overline{g}$ commutes with s. But

$$\begin{aligned}
(g\overline{g}) \star s &= g\overline{g}s && (\star \text{ used for emph}) \\
&= g(\overline{g}s) \\
&= g(s\overline{g}) && (\text{Since } \overline{g} \in C_G(S) \Rightarrow \overline{g}s = s\overline{g}) \\
&= (gs)\overline{g} \\
&= (sg)\overline{g} && (\text{Since } \overline{g} \in C_G(S) \Rightarrow gs = sg) \\
&= s \star (g\overline{g}).
\end{aligned}$$

TS2 Now we need to show that $g^{-1} \in C_G(s)$, i.e. that g^{-1} commutes with s whenever g does. This is proved as follows:

$$\begin{aligned}
gs = sg & && (\text{Since } g \in C_G(s)) \\
\Rightarrow gsg^{-1} = s & && (\text{Mul bs on right by } g^{-1}) \\
\Rightarrow sg^{-1} = g^{-1}s & && (\text{Mul bs on left by } g^{-1})
\end{aligned}$$

which, by definition means that g^{-1} commutes with s, i.e. $g^{-1} \in C_G(s)$.

∎

All this gives us a nice proof of the fact that if a and b commute then so does a^n and b^m for any integer n and m. Indeed,

$$\begin{aligned}
a, b \text{ commute} &\Rightarrow a \in C_G(b) \\
&\Rightarrow a^n \in C_G(b) && (C(b) \leqslant G) \\
&\Rightarrow b \in C_G(a^n) && (\text{Proposition 6.33}) \\
&\Rightarrow b^m \in C_G(a^n) && (C(a^n) \leqslant G) \\
&\Rightarrow a^n, b^m \text{ commute.}
\end{aligned}$$

This is the slick proof of Theorem 5.32**(4)** I had promised earlier.

EXERCISE 6.16. Show that for any group G, $\mathcal{C}(G)$ is abelian.

SOLUTION

EXERCISE 6.17. Show that a group is abelian iff $G = \mathcal{C}(G)$.

SOLUTION

This last exercise says that in some way the center of a group is a measure of how close a group is to being abelian.

By the very definition of the identity, we have that $e_G g = g e_G = g$ for every $g \in G$. Ignoring the last equality means that $ge = eg$, i.e that e_G and g commute. Hence the identity is in the center of any group; we say that the a group is *centerless* if its center contains only this. That is G is centerless iff $\mathcal{C}(G) = \{e_G\}$.

WARNING! 6.36. Saying that a group is centerless does not mean that the center of that group is empty! Again, the center of a group is **never** empty. It just means that it's as small as possible!

EXAMPLE 6.37. The table

$\mathcal{D}_4$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$
$\mathbf{1}$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$
ρ	ρ	ρ^2	ρ^3	$\mathbf{1}$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$
ρ^2	ρ^2	ρ^3	$\mathbf{1}$	ρ	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$
ρ^3	ρ^3	$\mathbf{1}$	ρ	ρ^2	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho$	$\mathbf{r}$
$\mathbf{r}$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{1}$	ρ	ρ^2	ρ^3
$\mathbf{r}\rho$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	ρ^3	$\mathbf{1}$	ρ	ρ^2
$\mathbf{r}\rho^2$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	ρ^2	ρ^3	$\mathbf{1}$	ρ
$\mathbf{r}\rho^3$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	ρ	ρ^2	ρ^3	$\mathbf{1}$

shows us that ρ^2 (which geometrically is a 180° degree rotation) is in the center of $\mathcal{D}_4$ (since reading the green row across is the same as reading the green column down). We also know that the identity is in the center. It's also easy to see that no other element is in the center of $\mathcal{D}_4$. Therefore, $\mathcal{C}(\mathcal{D}_4) = \{\mathbf{1}, \rho^2\}$. The table can also be checked to see that

$$C_{\mathcal{D}_4}(\rho) = \{\mathbf{1}, \rho, \rho^2, \rho^3\} = \langle \rho \rangle \quad \text{and} \quad C_{\mathcal{D}_4}(\mathbf{r}) = \{\mathbf{1}, \mathbf{r}, \mathbf{r}\rho^2, \rho^2\}.$$

$\mathcal{D}_4$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$
$\mathbf{1}$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$
ρ	ρ	ρ^2	ρ^3	$\mathbf{1}$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$
ρ^2	ρ^2	ρ^3	$\mathbf{1}$	ρ	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$
ρ^3	ρ^3	$\mathbf{1}$	ρ	ρ^2	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho$	$\mathbf{r}$
$\mathbf{r}$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{1}$	ρ	ρ^2	ρ^3
$\mathbf{r}\rho$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	ρ^3	$\mathbf{1}$	ρ	ρ^2
$\mathbf{r}\rho^2$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	ρ^2	ρ^3	$\mathbf{1}$	ρ
$\mathbf{r}\rho^3$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	ρ	ρ^2	ρ^3	$\mathbf{1}$

EXAMPLE 6.38. In practice to find the center of a group one takes the centralizer of a few elements and intersects them. Then we take some obvious elements in the center and tries to show that this intersection contains only these obvious elements. For instance, our goal is to find $\mathcal{C}(\mathbf{GL}_2(R))$, but we start off by finding the centralizers of the specific elements

$$\mathfrak{L} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathfrak{R} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

of $\mathbf{GL}_2(R)$ (and even $\mathbf{SL}_2(R)$). To this end, let $\mathbf{A}$ be the arbitrary invertible matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then we have that

$$\mathfrak{L} \cdot \mathbf{A} = \begin{pmatrix} a+c & b+d \\ c & d \end{pmatrix} \quad \text{and} \quad \mathbf{A} \cdot \mathfrak{L} = \begin{pmatrix} a & b+a \\ c & c+d \end{pmatrix}.$$

Thus if $\mathbf{A}$ is in the centralizer of $\mathfrak{L}$, these two matrices must be the same. So their difference must be zero.[II] But this difference is

$$\mathbf{A} \cdot \mathfrak{L} - \mathfrak{L} \cdot \mathbf{A} = \begin{pmatrix} c & d-a \\ 0 & c \end{pmatrix}$$

which implies that $c = 0$ and $a = d$. Conversely, following our logic backwards, any matrix satisfying these two properties must commute with $\mathfrak{L}$. Therefore we have that

$$C_{\mathbf{GL}_2(R)}(\mathfrak{L}) = \left\{ \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} : a \in \mathcal{U}(R), b \in R \right\}.$$

This also shows that[III]

$$C_{\mathbf{SL}_2(R)}(\mathfrak{L}) = \left\{ \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} : a^2 = 1, b \in R \right\}.$$

[II] Just because the group operation is matrix multiplication doesn't mean that we can't take advantage of the other goodies which come with matrices.

[III] For instance, for R a field, $\mathbb{F}$, (i.e. $\mathbb{F} = \mathbb{Z}_p$ (p-prime), $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$) we have that

$$C_{\mathbf{SL}_2(\mathbb{F})}(\mathfrak{L}) = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} : b \in \mathbb{R} \right\} \cup \left\{ \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix} : b \in \mathbb{F} \right\}.$$

At the same time, we could just as easily show [IV]

$$C_{\mathbf{GL}_2(R)}(\mathfrak{R}) = \left\{\begin{pmatrix} a & 0 \\ b & a \end{pmatrix} : a \in \mathcal{U}(R)\right\}. \quad \text{and} \quad C_{\mathbf{SL}_2(R)}(\mathfrak{R}) = \left\{\begin{pmatrix} a & b \\ 0 & a \end{pmatrix} : a^2 = 1, b \in R\right\}.$$

Therefore, the intersection of these two centralizers in $\mathbf{GL}_2(R)$ is the set of invertible, scalar matrices, which we've denoted as $\mathcal{S}\mathbf{GL}_2(R)$ and the intersection in $\mathbf{SL}_2(R)$ is the set $\mathcal{S}\mathbf{SL}_2(R)$ (So

$$\mathcal{S}\mathbf{GL}_2(R) = \left\{\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} : a \in \mathcal{U}(R)\right\} \quad \text{and} \quad \mathcal{S}\mathbf{GL}_2(R) = \left\{\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} : a^2 = 1\right\}.$$ [V]

) Now as an intersection of centralizers, $\mathcal{S}\mathbf{GL}_2(R) \preccurlyeq \mathcal{C}(\mathbf{GL}_2(R))$ and $\mathcal{S}\mathbf{SL}_2(R) \preccurlyeq \mathcal{C}(\mathbf{SL}_2(R))$. Conversely, it's easy to see that every element of $\mathcal{S}\mathbf{GL}_2(R)$ commutes with any matrix and so $\mathcal{C}(\mathbf{GL}_2(R)) \preccurlyeq \mathcal{S}\mathbf{GL}_2(R)$. Therefore, since we have proven that inclusion holds in both directions, equality must hold and

$$\mathcal{S}\mathbf{GL}_2(R) = \mathcal{C}(\mathbf{GL}_2(R)) \quad \text{and} \quad \mathcal{S}\mathbf{SL}_2(R) = \mathcal{C}(\mathbf{SL}_2(R)).$$

EXAMPLE 6.39. In this example we look at the centers and centralizers of groups associated with quaternions.

(1) Let $\mathsf{q} \in \mathbb{H}^\times$ be non-real. That is

$$\mathsf{q} = \mathsf{q}_R + \mathbf{q}_P, \ a \in \mathbb{R}, \mathbf{q}_P \neq 0 \in \mathbb{R}^3.$$

Then since

$$\mathsf{q} * \mathsf{p} - \mathsf{p} * \mathsf{q} = 2 = 2\,\mathbf{q}_P \times \mathbf{p}_P$$

(where $\mathbf{p}_P$ is the pure part of p) we have that p commutes with q iff the pure part of p is proportional to the pure part of q. The set of all such p was previously denoted as $\mathbb{C}_{\mathbf{v}}$. (Recall this was a copy of the complex numbers where $\frac{\mathbf{q}_P}{\|\mathbf{q}_P\|}$ played the role of $\mathbf{i}$.) Thus

$$C_{\mathbb{H}^\times}(\mathsf{q}) = \mathbb{C}^\times_{\mathbf{q}_P}.$$ [VI]

Let's ask our selves the same questions for unit quaternions, i.e. elements of $\mathbb{S}^3$. Again, if $\mathsf{q} = \pm 1$, the centralizer of q is everything, in this case $\mathbb{S}^3$. All other quaternions in $\mathbb{S}^3$ have a non-zero pure part. Let $\mathsf{q} \in \mathbb{S}^3$ be such a quaternion. Then we have

$$\mathsf{q} = \cos\theta + \sin\theta\,\widehat{\mathbf{u}}$$

for some $\widehat{\mathbf{u}} \in \mathbb{S}^2$ and $0 < \theta < 2\pi$. And the only elements of $\mathbb{S}^3$ which commute with q are elements who have a similar factorization with the same $\widehat{\mathbf{u}}$ (and with a possibly different angle instead of θ). In our previous notation

$$C_{\mathbb{S}^3}(\mathsf{q}) = \mathbb{S}^1_{\widehat{\mathbf{u}}}$$

or in more detail

$$C_{\mathbb{S}^3}(\cos\theta + \sin\theta\,\widehat{\mathbf{u}}) = \{\cos\psi + \sin\psi\,\widehat{\mathbf{u}} : 0 \leq \psi < 2\pi\}.$$

For example, if we simply consider $\mathbf{i}$ and $\mathbf{j}$ we see that

$$C_{\mathbb{H}^\times}(\mathbf{i}) \cap C_{\mathbb{H}^\times}(\mathbf{j}) = \mathbb{C}^\times_{\mathbf{i}} \cap \mathbb{C}^\times_{\mathbf{j}} = \mathbb{R}^\times$$

[IV] Actually, using the operating reversing nature of the transpose, it's easy to see that matrices $\mathbf{A}$ and $\mathbf{B}$ commute iff their transposes, $\mathbf{A}^\top$ and $\mathbf{B}^\top$ do. As such, since $\mathfrak{L}$ and $\mathfrak{R}$ are transposes of each each other, the centralizer of the one is the (set of) transpose of the other.

[V] If R were a field, $\mathcal{S}\mathbf{SL}_2(R) = \pm\mathbf{I}_2$.

[VI] Recall $\mathbb{C}_{\mathbf{q}_P}$ is by definition the set of all non-zero quaternions whose pure part is proportional to $\mathbf{q}_P$.

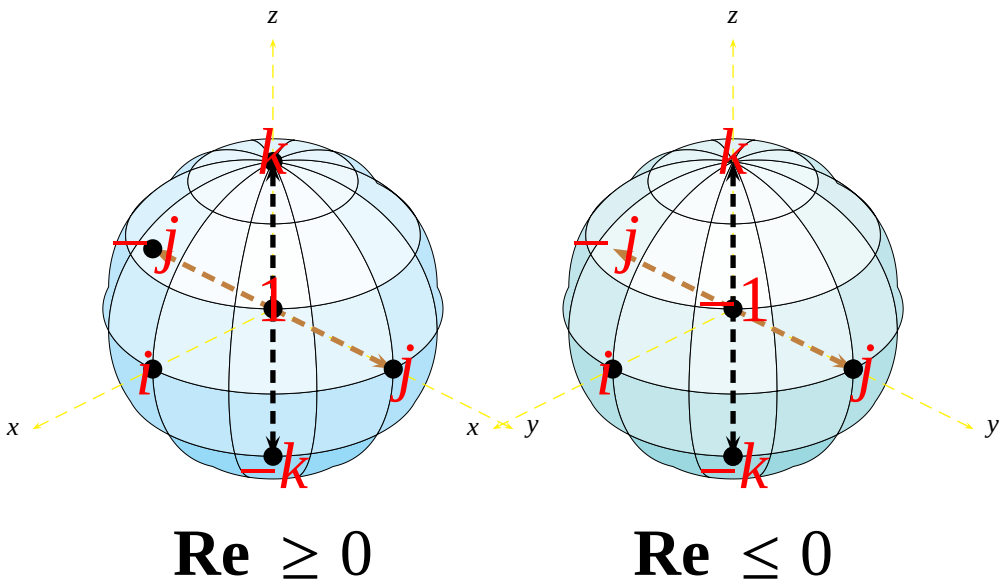


Figure 6.1: The centralizer of **j** is highlighted in **brown**, and the centralizer of **k** is highlighted in **black**. Note they only meet at ±1

and

$$C_{\mathbb{S}^3}(\mathbf{i}) \cap C_{\mathbb{S}^3}(\mathbf{j}) = \mathbb{S}^1_{\mathbf{i}} \cap \mathbb{S}^1_{\mathbf{j}} = \{\pm 1\}.$$

Therefore we have computed the centers of the groups $\mathbb{H}^\times$ and $\mathbb{S}^3$:

$$\mathcal{C}(\mathbb{S}^3) = \{\pm 1\} \quad \text{and} \quad \mathcal{C}(\mathbb{H}^\times) = \mathbb{R}^\times.$$

(2) Looking instead at the double dihedral groups, $B\mathcal{D}_n$, then we have, by definition both $\mathbf{j} \in B\mathcal{D}_n$ as well as some element $\mathbb{S}^1_{\mathbf{i}}$ (the element χ_{2n} is in there). Thus, in this case, we also have that

$$\mathcal{C}(B\mathcal{D}_n) = \{\pm 1\} \text{ for every } n.$$

(3) More generally, if $H \preccurlyeq \mathbb{S}^3$ and it contains two quaternions with non-paralled axis, then $\mathcal{C}(H)$ is equal to $\{1, -1\}$ if -1 is in H and is simply equal to just 1 if not. (i.e. the center is $\{1, -1\} \cap H$).

(4) The next example uses material from Appendix C, Reflections and Rotations. If we pass to $\mathbf{SO}_3(\mathbb{R})$, the group of rotations in three dimensions, then the centralizer of $\mathbf{R}_\mathsf{q}$ (where q is the same as above) would be described as (the image of) all the quaternions from $\mathbb{S}^1_{\widehat{\mathbf{u}}}$ (i.e. the set of rotations with the same axis as q). However, since $\mathbf{R}_{-1} = \mathbf{R}_1 = \mathbf{1}_{\mathbb{R}^3}$ if $\mathsf{qp} = -\mathsf{pq}$, then the since

$$\mathbf{R}_\mathsf{q} \circ \mathbf{R}_\mathsf{p} = \mathbf{R}_{\mathsf{qp}} = \mathbf{R}_{-\mathsf{pq}} = \mathbf{R}_{-1} \circ \mathbf{R}_\mathsf{p} \circ \mathbf{R}_\mathsf{q} = \mathbf{R}_\mathsf{p} \circ \mathbf{R}_\mathsf{q}$$

we'd also have that the corresponding rotations would also commute. In Theorem 3.93 back on 118, we showed this is possible iff q and p are orthogonal, pure quaternions. In conclusion, if a non-trivial rotation is by an angle less than 180 degrees, then the only rotations which commute with it are the rotations which have the same axis. On the other hand, if the rotation is by 180 degrees, then in addition to commuting with rotations which share the original's axis, it also commutes with all other 180° rotations about axis which are orthogonal to the original's axis. Therefore there are three cases:

(a) Of course rotation by a multiple of 360° is the identity, and as such, commutes with everything.

(b) If $\rho \in \mathbf{SO}_3(\mathbb{R})$ is a rotation by an angle which is not a multiple of 180°, then the centralizer of ρ consists only those rotations which have the same axis as ρ.

(c) If $\rho \in \mathbf{SO}_3(\mathbb{R})$ is a rotation by 180°, then in addition to commuting with all rotations which have the same axis as ρ does, ρ also commutes with all 180° rotations about any axis which is perpendicular to ρ's. For instance, the rotations $\varrho_{\pi,\mathbf{i}}$ and $\varrho_{\pi,\mathbf{j}}$ commute.

Now we can easily prove that $\mathbf{SO}_3(\mathbb{R})$ is centerless (i.e. its center contains only the identity.) For instance, if $\mathbf{v}$ and $\mathbf{u}$ determine distinct axis (i.e. are not parallel), then for any angles ϕ and θ less than π, we would have that only the identity commutes with both $\varrho_{\phi,\mathbf{u}}$ and $\varrho_{\theta,\mathbf{v}}$. For instance, only the identity can commute with both $\varrho_{\mathbf{i},\pi/2}$ and $\varrho_{\mathbf{j},\pi/2}$. There we have proven that $\mathbf{SO}_3(\mathbb{R})$ is centerless, i.e.

$$\mathcal{C}(\mathbf{SO}_3(\mathbb{R})) = \{\mathbf{1}_{\mathbb{R}^3}\}.$$

Exercise 6.18. List all the elements in

(1) $C_{\mathbf{SO}_3(\mathbb{R})}(\varrho_{\pi,\mathbf{i}}) \cap C_{\mathbf{SO}_3(\mathbb{R})}(\varrho_{\pi,\mathbf{j}})$,

(2) $C_{\mathbf{SO}_3(\mathbb{R})}(\varrho_{\pi,\mathbf{i}+\mathbf{j}}) \cap C_{\mathbf{SO}_3(\mathbb{R})}(\varrho_{\pi,\mathbf{j}})$,

(3) $C_{\mathbf{SO}_3(\mathbb{R})}(\varrho_{\pi/2,\mathbf{i}}) \cap C_{\mathbf{SO}_3(\mathbb{R})}(\varrho_{\pi,\mathbf{j}})$,

(4) $C_{\mathbf{SO}_3(\mathbb{R})}(\varrho_{\pi/2,\mathbf{i}}) \cap C_{\mathbf{SO}_3(\mathbb{R})}(\varrho_{\pi/2,\mathbf{j}})$.

SOLUTION

Exercise 6.19. For any subset $S \subseteq G$ show that

$$C_G(S) = \{g \in G : \forall s \in S,\ gsg^{-1} = s\}.$$

(This is a one line proof!)

SOLUTION

Exercise 6.20.

(1) Let $G \leqslant \mathbb{S}^1$ be a finite group of unit quaternions. Prove that $G = \{\{n\}\}$ for some positive integer n (in particular, every finite subgroup of $\mathbb{S}^1$ is cyclic.)
Hint! Consider the element of G which has the smallest angle.

(2) Let $B \leqslant \mathbb{S}^3$ be finite and $\mathsf{q} \in B$ have a non-zero pure part. Prove that the centralizer of q is cyclic.

(3) Prove that any non-cyclic subgroup of $\mathbb{S}^3$ has a center contained inside of $\{\pm 1\}$.
Hint! Show that any abelian subgroup of $\mathbb{H}^\times$ is contained inside of a longitude:

$$\mathbb{S}^1_{\widehat{\mathbf{u}}} = \{\cos(\theta) + \sin(\theta)\,\widehat{\mathbf{u}} : \theta \in \mathbb{R}\}$$

for some $\widehat{\mathbf{u}} \in \mathbb{S}^2$. Use the fact that these sets behave exactly the same as $\mathbb{S}^1$ and the previous part.

SOLUTION

If $g \in \mathcal{C}(G)$, then, by definition of the center of a group, everybody in G commutes with g. Thus $C_G(g) = G$. However G needn't be abelian to have non-trivlal central elements.

Exercise 6.21. Give an example of a group G and a non-central element of G whose centralizer is non-abelian.
Hint! The previous exercise shows that $\mathbb{S}^3$ cannot be a source for examples of this.

Solution

Exercise 6.22. Find $\mathcal{C}(\mathrm{Af}_p)$ where Af_p is the group from Exercise 6.6 on page 323.

Solution

Exercise 6.23. Show that $\mathbf{SL}_3(\mathbb{Z}_2)$ is centerless.
Hint! Again, don't write down what it means for two arbitrary matrices in this group to commute, but pick out some simple matrices and figure out what it means for all matrices to commute with the ones you've picked. In this way you end up considering a couple of linear problems instead of a single (very difficult) non-linear problem.

Solution

Exercise 6.24.

Solution

(1) Let R be a commutative ring with identity (again, for our purposes this means that $R = \mathbb{Z}_n, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$, or $\mathbb{C}$.) Show that the matrices

$$\left\{\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} : a, b, c \in R\right\}$$

form a group by showing that they form a subgroup of $\mathbf{SL}_3(R)$. This group is called the Heisenberg group (after the physicist) and is denoted $\mathrm{UT}_3(R)$ (UT stands for *upper triangular*). In the case that $R = \mathbb{Z}_n$ show that this group has order n^3. (We have already shown that every non-identity element of $\mathrm{UT}_3\left(\mathbb{Z}_p\right)$ where p is an odd prime has order p).

(2) Show that the center of $\mathrm{UT}_3(R)$, $\mathcal{C}(\mathrm{UT}_3(R))$ is

$$\left\{\begin{pmatrix} 1 & 0 & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} : b \in R\right\}.$$

In particular, these groups are never abelian expect in the stupid case that R contains exactly 1 element. One possible outline for this is as follows:

- For the time being, call the proposed center H. Show that if an arbitrary matrix

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$$

in $\mathrm{UT}_3(R)$ commutes with the matrix

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

then $c = 0$.

- Show that if that same arbitrary matrix commutes with

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

then we must have that $a = 0$ as well.

- Conclude that $\mathcal{C}(\mathrm{UT}_3(R)) \subseteq H$.
- Conversely, show that every element of H commutes with an arbitrary element of $\mathrm{UT}_3(R)$. Thus we have that $H \subseteq \mathcal{C}(\mathrm{UT}_3(R))$.
  Careful! Be careful not to use multiple meanings for the letter b while doing this!

Exercise 6.25. Let **C** denote the cat matrix

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Solution

Calculate the centralizer of **C** in

(1) $\mathbf{SL}_2(\mathbb{Z}_3)$ **(2)** $\mathbf{SL}_2(\mathbb{Z}_{11})$

by showing that both are cyclic. Prove in the first case that this is generated by **C** and in the second case, it cannot be generated by **C**.[VII]
Hint! This is harder than the other matrix examples. It comes down to solving the equation

$$x^2 + xy - y^2 = 1$$

over $\mathbb{Z}_3$ and then $\mathbb{Z}_{11}$. As of right now we don't know a trick to do this, so you will probably have to use brute force. Calculate the order of **C** in both cases. Show that for $\mathbb{Z}_3$ the number of solutions is the same as the order and for $\mathbb{Z}_{11}$ it is not.

Exercise 6.26. Let $\mathbf{L}_{13}$ be the 3 by 3 matrix

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Solution

regarded as a matrix in $\mathbf{M}_3(\mathbb{Z}_p)$.

(1) Prove that $C_{\mathbf{GL}_3(\mathbb{Z}_p)}(\mathbf{L}_{13})$ is

$$\left\{\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & a \end{pmatrix} : a, d \neq 0\right\}.$$

(2) Show that for every prime p, this is a non abelian (hence non–cyclic) subgroup of order $(p-1)^2 \cdot p^3$.

(3) Prove $C_{\mathbf{SL}_3(\mathbb{Z}_p)}(\mathbf{L}_{13})$ is still non abelian (for every prime), but has order $(p-1) \cdot p^3$.

Let $c \in \mathbb{Z}_p^\times$, so that c^{-1} is defined. Define

$$\mathbf{D}_c := \begin{pmatrix} c & 0 \\ 0 & c^{-1} \end{pmatrix}.$$

Note that $\mathbf{D}_c \in \mathbf{SL}_2\left(\mathbb{Z}_p\right)$ while $\mathbf{D}_1 = \mathbf{I}_2$ and $\mathbf{D}_{-1} = -\mathbf{I}_2$. Note also that for any $c, c' \in \mathbb{Z}_p^\times$ we have

$$\mathbf{D}_c \cdot \mathbf{D}_{c'} = \mathbf{D}_{c \cdot c'} = \mathbf{D}_{c'} \cdot \mathbf{D}_c.$$

In particular

$$\{\mathbf{D}_{c'} : c' \in \mathbb{Z}_p^\times\} \subseteq C_{\mathbf{SL}_2(\mathbb{Z}_p)}\left(\mathbf{D}_c\right).$$

For most $c \in \mathbb{Z}_p^\times$ this ends up being the entire centralizer. That is the content of the following Example/Theorem:

Example /Theorem 6.40. Let $c \neq \pm 1$ be an element of $\mathbb{Z}_p$, p a prime, and consider the matrix. Then

$$C_{\mathbf{SL}_2(\mathbb{Z}_p)}\left(\mathbf{D}_c\right) = \{\mathbf{D}_a : a \in \mathbb{Z}_p^\times\}.$$

Remark 6.41. In other words,

$$C_{\mathbf{SL}_2(\mathbb{Z}_p)}\left(\mathbf{D}_c\right) = \left\{\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} : a \in \mathbb{Z}_p^\times\right\}.$$

In particular, since $\mathbb{Z}_p^\times = \mathcal{U}(\mathbb{Z}_p)$ is cyclic, it has a generator, call it β, $C_{\mathbf{SL}_2(\mathbb{Z}_p)}\left(\mathbf{D}_c\right)$ is cyclic with generator

$$\mathbf{D}_\beta = \begin{pmatrix} \beta & 0 \\ 0 & \beta^{-1} \end{pmatrix}.$$

Of course if we allowed c to equal 1 or -1, then $\mathbf{D}_c$ would equal either $\mathbf{I}_2$ or $-\mathbf{I}_2$, and the statement of the Example/Theorem would be false as $C_{\mathbf{SL}_2(\mathbb{Z}_p)}\left(\mathbf{D}_c\right)$ would wind up being all of $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$.

Proof. Let $\mathbf{B}$ be an arbitrary matrix

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

in $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ with commutes with $\mathbf{D}_c$. A direct calculation shows that we have that

$$\mathbf{D}_c\mathbf{B}\mathbf{D}_c^{-1} = \begin{pmatrix} b_{11} & c^2 b_{12} \\ \frac{1}{c^2} b_{21} & b_{22} \end{pmatrix}.$$

But since $\mathbf{D}_c$ and $\mathbf{B}$ are assumed to commute, we must have that

$$\mathbf{C}_c\ \mathbf{B}\mathbf{D}_c^{-1} = \mathbf{B}.$$

But this implies that

$$c^2 \cdot b_{12} = b_{12} \text{ and } \frac{1}{c^2} \cdot b_{21} = b_{21}.$$

Now if either b_{12} or b_{21} were non-zero, we could cancel it and we'd be able to conclude that $c^2 = 1$. But in $\mathbb{Z}_p$, this would imply that $c = 1$ or $c = -1$, a contradiction. Therefore we indeed must have both non-diagonal entries, b_{12} and b_{21} are zero and so

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} b_{11} & 0 \\ 0 & b_{22} \end{pmatrix}.$$

Lastly note that $b_{11}^{-1} = b_{22}$ since we are also assuming that $\mathbf{B} \in \mathbf{SL}_2\left(\mathbb{Z}_p\right)$. In conclusion, $\mathbf{B} = \mathbf{D}_{b_{11}}$, which completes the proof. ■

See Definition 3.237 back on page 211 for the definitions in this exercise.

EXERCISE 6.27. Let $\mathbb{F}$ be a field.

SOLUTION

(1) (You only need to do this part if you cannot do the second part, but this first part may be a necessary warmup for you) Let $\mathbf{d} := (d_1, \ldots, d_n) \in (\mathbb{F}^\times)^{*n}$ (i.e. the d_i are n *distinct* non-zero elements of $\mathbb{F}$). Let $\mathbf{D}(\mathbf{d})$ be the diagonal matrix. Prove that

$$C_{\mathbf{GL}_n(\mathbb{F})}(\mathbf{D}(\mathbf{d})) = \mathrm{G}\delta(n, \mathbb{F})$$

(2) Now we generalize the previous part to the case that entries of $\mathbf{d}$ are not necessarily distinct (but still non-zero): say we have

$$\mathbf{d} = (\underbrace{d_1, \ldots, d_1}_{n_1\text{-times}}, \underbrace{d_2, \ldots, d_2}_{n_2\text{-times}}, \ldots, \underbrace{d_k, \ldots, d_k}_{n_k\text{-times}}).$$

Prove that

$$C_{\mathbf{GL}_n(\mathbb{F})}(\mathbf{D}(\mathbf{d})) = \mathbf{GL}_{n_1}(\mathbb{F}) \oplus \mathbf{GL}_{n_2}(\mathbb{F}) \oplus \cdots \oplus \mathbf{GL}_{n_k}(\mathbb{F}).$$

Clarification! Part of the problem is to figure out what the notation

$$\mathbf{GL}_{n_1}(\mathbb{F}) \oplus \mathbf{GL}_{n_2}(\mathbb{F}) \oplus \cdots \oplus \mathbf{GL}_{n_k}(\mathbb{F})$$

means. What you should figure out what the centralizer is and then say to yourself, oh, that's what this terrible author crummy notation must mean.

EXAMPLE /THEOREM 6.42. Suppose that $|g| = 4n$ and $z \in C(g)$ has order 2. Then $|zg| = |g|$.

PROOF. For suggestivity, we denote the element z as (-1) so that zg is $(-1) \cdot g$.[VIII] We also choose to denote the identity as 1. Let m denote the order of $|(-1) \cdot g|$. Since $|g|$ is even, $((-1) \cdot g)^{|g|} = (-1)^{|g|} \cdot g^{|g|} = 1 \cdot 1 = 1$, we know that $m \Big| |g|$. Moreover, since

$$g^{2m} = (-1)^{2m} \cdot g^{2m} = ((-1) \cdot g)^{2m} = (((-1) \cdot g)^m)^2 = 1^2 = 1$$

we also know that in addition $|g| \Big| 2m$. These two facts clearly imply that m either equals $|g|$ or it equals $\frac{|g|}{2}$. We now show the second case is impossible. As $|g|$ is a multiple of 4, $\frac{|g|}{2}$ is even and so

$$((-1) \cdot g)^{\frac{|g|}{2}} = (-1)^{\frac{|g|}{2}} \cdot g^{\frac{|g|}{2}} = 1 \cdot g^{\frac{|g|}{2}} = g^{\frac{|g|}{2}}.$$

But since $\frac{|g|}{2} < |g|$, we cannot possibly have that $g^{\frac{|g|}{2}} = 1$. Therefore, the only possible case which remains is $m = |g|$, as desired. ∎

6.6 SMALLEST SUBGROUP CONTAINING TWO SUBGROUPS

Let H and K be subgroups of a group G. Obviously the largest subgroup contained H and K is the intersection of H and K. This means that if $N \leqslant H$ and $N \leqslant K$ holds iff $N \leqslant H \cap K$. We would naively hope that, just like sets in general, the opposite property would be held by the union of H and K, $H \cup K$. However the union of two subgroups is rarely every a subgroup as the following exercise illustrates:

EXERCISE 6.28. Suppose that $H, K \leqslant G$. Show that $H \cup K \leqslant G$ iff $H \subseteq K$ or $K \subseteq H$.

SOLUTION

This surprises most students. In many cases the following definition "plays the role of" $H \cup K$. We define it for arbitrary subsets of a group G but as you should guess we will only really be interested in this concept in the case that H and K are subgroups (and not just any old subset) of G.

The following definition is a direct generalization of the definition given for the integers, 3.5 back on page 74.

DEFINITION 6.43. Suppose that H and K are subsets of a group G with binary operation $\star$. Then

$$H \star K \stackrel{\text{def}}{=} \{h \star k : h \in H, k \in K\}.$$

Note that an element $g \in G$ is in $H \star K$ iff it can, in some sense, be factored as element of H times an element of K. That is iff there exists $h \in H$ and $k \in K$ so that $g = hk$. As you should expect, if the group G is written multiplicatively (which we generically always assume to be the case), then we write this simply as HK and if G is written additively then we write this as $H + K$.

[VIII] This doesn't mean we can start adding things together!

Example /Theorem 6.44. For any three subsets H, K, and N of a group G we have that

$$(HK)\,N = H\,(KN)\,.$$

Proof. Let $g \in (HK)\,N$ be arbitrary. Then there exists $g_1 \in H\,K$ and $n \in N$ so that $g = g_1 n$. But since $g_1 \in H\,K$ there exists $h \in H$ and $k \in K$ so that $h_1 = hk$. But this implies that $g = (hk)n = h(kn)$. But this latter product is in $H\,(KN)$. This proves that

$$(HK)\,N \subseteq H\,(KN)\,.$$

The proof of the reverse inclusion is the same. ∎

Remark 6.45. It is not true in general that $H\,K$ is a subgroup of G even if both H and K are. For instance, we take

$$G = \mathcal{D}_4, \qquad H = \langle\, \mathbf{r} \,\rangle = \{\mathbf{1}, \mathbf{r}\} \quad \text{and} \quad K = \langle\, \mathbf{r}\rho \,\rangle = \{\mathbf{1}, \mathbf{r}\rho\}.$$

Then

$$\begin{aligned} H\,K &= \{\mathbf{1}, \mathbf{r}\} \cdot \{\mathbf{1}, \mathbf{r}\rho\} \\ &= \{\mathbf{11}, \mathbf{1r}\rho, \mathbf{r1}, \mathbf{rr}\rho\} \\ &= \{\mathbf{1}, \mathbf{r}\rho, \mathbf{r}, \rho\}. \end{aligned}$$

But this is not a subgroup of $\mathcal{D}_4$ since it is not closed under multiplication:

$$\rho \in H\,K \text{ but } \rho \circ \rho = \rho^2 \notin H\,K.$$

We now give a sufficient condition which guarantees that $H\,K \leqslant G$.

Definition 6.46. Let $H, K \leqslant G$. We say that K *normalizes* H if for all $k \in K$ and $h \in H$ we have that $khk^{-1} \in H$.

For instance, by closure, H always normalizes itself. Moreover if G is an abelian group, then any two subgroups normalize each other. This is because in this case, if $h \in H$ and $k \in K$, then since G is abelian we have

$$\begin{aligned} khk^{-1} &= kk^{-1}h \quad (G \text{ assumed abelian}) \\ &= h. \end{aligned}$$

For the remainder of this chapter we are most interested in this abelian case. However later we will be least interested in that possibility!

As of right now we have no other real choice of how to calculate this besides the old fashioned straight up way. Here's a "fun" example of this ... joy:

Example 6.47. Let $H = \langle \mathbf{i} \rangle \leqslant \mathbb{H}^\times$ and $K = \langle \mathbf{j} \rangle \leqslant \mathbb{H}^\times$. Then

$$\begin{aligned} HK &= \{1, \mathbf{i}, -1, -\mathbf{i}\} \cdot \{1, \mathbf{j}, -1, -\mathbf{j}\} \\ &= \{1, \mathbf{j} - 1, -\mathbf{j}\} \cup \{\mathbf{i}, \mathbf{k}, -\mathbf{i}, -\mathbf{k}\} \cup \{-1, -\mathbf{j}, 1, \mathbf{j}\} \cup \{-\mathbf{i}, -\mathbf{k}, \mathbf{i}, \mathbf{k}\} \\ &= \{1, -1, \mathbf{i}, -\mathbf{i}, \mathbf{j}, -\mathbf{j}, \mathbf{k}, -\mathbf{k}\} \\ &= Q_8. \end{aligned}$$

The following theorem says that when one of the two subgroups involved normalizes the other, we do indeed have that HK is a subgroup of G. Moreover, you should think of the other parts as saying that it is the smallest subgroup of G containing both H and K. This is the first time in these notes that we have made active use of a normal like hypothesis, and as such you should pay extra attention to how it is used.

Theorem 6.48. *Suppose that H and K are subgroups of a group G. Then*

(1) $H \cup K \subseteq HK$

(2) $H \cup K \subseteq N \leqslant G \Rightarrow HK \subseteq N$,[I].

Now suppose that either H normalizes K or that K normalizes H. Then with this additional hypothesis we have that

(3) $HK \leqslant G$.

Proof.

TS1 Let $h \in H$ be arbitrary. Since $e_G \in K$ we know that $he_G \in HK$. But this product is just h and therefore $H \subseteq HK$. The proof that $K \subseteq HK$ is the same.

TS2 Let $hk \in HK$. We need to show that $hk \in N$. But since N contains H we know that $h \in N$ and since N contains K we know that $k \in N$ as well. But since we are assuming that N is a group, we know that $hk \in N$ by the closure on N. Hence $HK \subseteq N$.

TS3 We can assume that K normalizes H. The proof when H normalizes K is essentially the same (after one or two cosmetic changes). Since $e_G \in H$ and K we have that $e_G = e_G * e_G \in HK$ which implies that HK is non-empty. We can therefore apply the two step subgroup test.

TS1 Now suppose that hk and $\dot{h}\dot{k}$ are in HK. For starters, since K normalizes H we know that

$$k\dot{h}k^{-1} = h' \in H. \tag{3}$$

Then we have that

$$\begin{aligned} hk \star \dot{h}\dot{k} &= hk\dot{h}\dot{k} && \text{(Only using } \star \text{ for emphasis)} \\ &= hk\dot{h}e\dot{k} \\ &= hk\dot{h}k^{-1}k\dot{k} \\ &= hh'k\dot{k} && \text{(Plugging in via (3))} \end{aligned}$$

Now $hh' = h'' \in H$ and $k\dot{k} \in K$ by closure, so the last product equals

$$\begin{aligned} &= h''k'' \\ &\in HK && \text{(By def'n of } HK\text{).} \end{aligned}$$

This proves closure under the operation.

[I] This should be compared with a defining property of the union of two sets. By this I mean, if A and B are two subsets of a set X, then $A \cup B$ is a subset of X and if A and B are both contained in a third subset of X, Y, then $A \cup B$ is also a subset of Y

TS2 Now let $hk \in HK$ be arbitrary[II]. Then

$$(hk)^{-1} = k^{-1}h^{-1}$$

by socks and shoes (Exercise 4.15). Now since K normalizes H we have that

$$k^{-1}h^{-1}k = \overline{h} \in H. \tag{4}$$

Therefore, we have that

$$k^{-1}h^{-1} = k^{-1}h^{-1}kk^{-1} = \overline{h}k^{-1} \in HK.$$

Therefore HK is closed under the taking of inverses.

Therefore HK passes the two step subgroup test and is therefore a subgroup of G. ■

It turns out that there are cases that $HK \leqslant G$ even if neither H nor G normalize each other. For instance there is a simple example of this in $\mathcal{S}_4$. In this exercise we refine the condition that $HK \leqslant G$.

EXERCISE 6.29. Suppose that $H, K \leqslant G$

(1) Prove that if $HK \leqslant G$, then $HK = KH$.
Hint! Let $h \in H$ and $k \in K$. Then $he_K \in HK$ and $e_Hk \in HK$. If $HK \leqslant G$, then use closure to conclude that $kh \in HK$.

(2) Prove that the converse of the previous part is also true. That is prove that if

$$HK = KH \Rightarrow HK \leqslant G.$$

Hint! Here is the proof of the closure of HK under inversion

PROOF. Assume that $HK = KH$. Let $hk \in HK$. Then by socks and shoes (Exercise 4.15)

$$(hk)^{-1} = k^{-1}h^{-1}.$$

Since $H \leqslant G$ H is closed under inversion. Hence $h^{-1} \in H$. Similarly $k^{-1} \in K$. Hence $k^{-1}h^{-1} \in KH$. But our assumption is that $KH = HK$. Hence $k^{-1}h^{-1} \in HK$ as well. Therefore $(hk)^{-1} \in HK$ whenever $hk \in HK$ and HK is therefore closed under inversion. ■

The proof that HK is also closed under multiplication is only slightly more difficult.

SOLUTION

Let's record the result from the previous exercise as a theorem:

THEOREM 6.49. *If $H, K \leqslant G$, then $HK \leqslant G$ iff the set HK is equal to the set KH.* ■

[II] As usual, we could put an extra step into this. We could say let $g \in HK$ be arbitrary. Then there exists $h \in H$ and $k \in K$ s.t. $g = hk$, but obviously this is not done in practice.

7 Symmetric Groups

Convention 7.1. For this chapter, we redefine what we mean by a mod n. We now require that this number live in $\overline{n} = \{1, \ldots, n\}$ instead of $\{0, \ldots, n-1\}$. Thus, whereas before 12 mod 12 was equal to the number 0, we now take it to be equal to 12. As such the only difference comes when we are looking at numbers which are divisible by n. [I]

Before getting started, we also throw in a definition that we will use throughout this chapter.

Definition 7.2. For a permutation $\sigma \in \mathcal{S}_n$, the *support* of σ is

$$\{i \in \overline{n} : \sigma[i] \neq i\}.$$ [II]

Elements of the support of σ are said to be *moved* by σ. The fixed point set of σ is defined to be

$$\{i \in \overline{n} : \sigma[i] = i\}.$$

Such letters are said to be *fixed* by σ.

Therefore the support and the fixed point set of any permutation partition $\overline{n}$ (that is they are disjoint and have a union equal to all of $\overline{n}$, i.e. every letter is in 1 and only 1 of them for a given σ).

7.1 Cycle Notation

So far we have been using the notation from Section 2.3.1 to describe a permutation. While this notation is a highly intuitive notation there is a far better notation for the purposes of group theory. For instance, to see what the order of a permutation is in the old notation is not a lot of fun. Also, a calculation is needed to figure out a permutation's parity. With the notation we develop here we will be able to calculate these two numbers essentially on sight. (Well, as easily as we can compute the least common multiple of a list of numbers).

Let $\sigma \in \mathcal{S}_n$. Like any function, σ is a relation[I] Since σ is a bijection, the corresponding digraph of σ consists of circles. This is because each of the finite number of vertices has exactly one arrow leaving it (that arrow is going to the unique image of the letter) and one arrow coming to it (which corresponds to the unique letter which

[I] This is similar to the annoying thing we do with clocks. Why the heck is 12 o'clock the first hour of the day? Shouldn't it be 0 o'clock or perhaps 1 o'clock?

[I] More explicitly we define a relation, also denoted σ, on the set $\overline{n}$ given by $i \sigma j$ if and only if $\sigma[i] = j$.

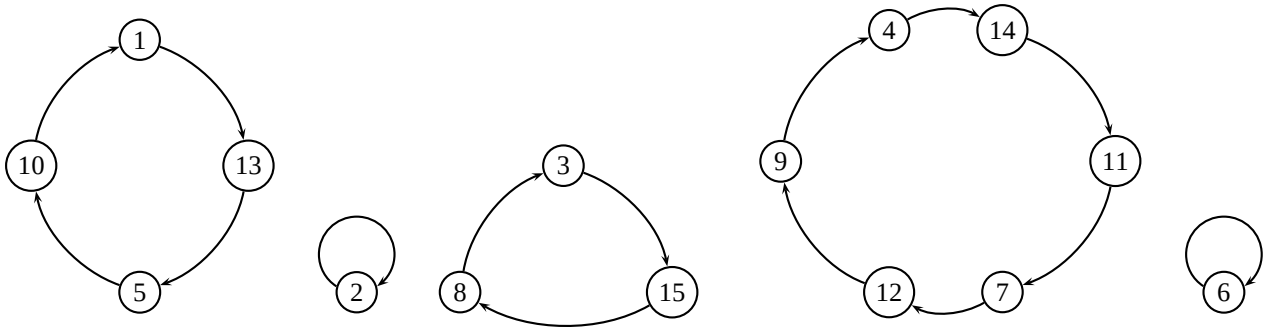
maps to the given letter). The circles are called the *cycles* of σ, and together these circles are called the *cycle decomposition* of σ. Notice that we can reverse this process: by taking a finite digraph consisting of circles[II] and turn it into a permutation. The corresponding permutation would take i to the unique vertex on the other end of the arrow which starts at i.

EXAMPLE 7.3.

(1) Suppose that σ is the permutation

$$\begin{array}{lllll} 1 \mapsto 13 & 2 \mapsto 2 & 3 \mapsto 15 & 4 \mapsto 14 & 5 \mapsto 10 \\ 6 \mapsto 6 & 7 \mapsto 12 & 8 \mapsto 3 & 9 \mapsto 4 & 10 \mapsto 1 \\ 11 \mapsto 7 & 12 \mapsto 9 & 13 \mapsto 5 & 14 \mapsto 11 & 15 \mapsto 8. \end{array}$$

Then the cycle decomposition is given by

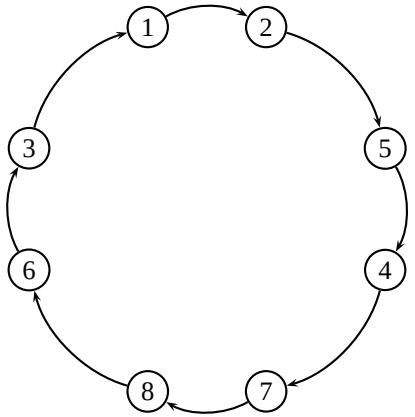


Notice the fact that the fixed point set of this permutation is $\{2, 6\}$ is reflected in the digraph by the fact that each have only an arrow leaving and coming right back to themselves.

(2) Suppose that σ is the permutation

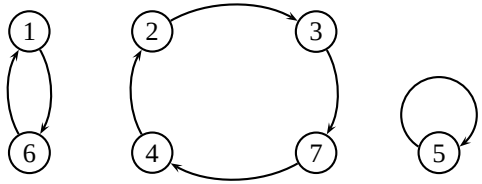
$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 5 & 1 & 7 & 4 & 3 & 8 & 6 \end{pmatrix}$$

Then the cycle decomposition is given by just the single circle:



[II] We do not consider a single vertex without any arrows going from it or to it a circle.

(3) Consider the digraph

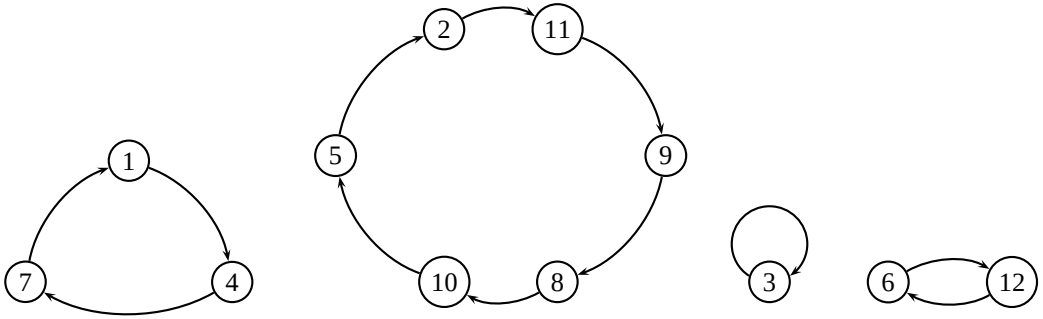


which consists only of circles. It corresponds to the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 6 & 3 & 7 & 2 & 5 & 1 & 4 \end{pmatrix}.$$

(4) Now let's consider the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 4 & 11 & 3 & 7 & 2 & 12 & 1 & 10 & 8 & 5 & 9 & 6 \end{pmatrix}$$



These circles aren't very compact, they can't really even fit on one line, so we use what is known as cycle notation to describe a cycle. For the following definition recall that

$$\overline{n}^{*m}$$

denotes the set of m tuples of positive integers between 1 and n with no two integers in the tuple the same.

DEFINITION 7.4. Suppose $m \leq n$, $(a_1, \ldots, a_m) \in \overline{n}^{*m}$. Then

$$(a_1\ a_2 \cdots a_m)$$

denotes the permutation (function) defined by

$$a \mapsto \begin{cases} a_{i+1} & \text{if } a = a_i \\ a & \text{if } a \neq a_i,\ i = 1 \ldots m \end{cases}$$

where all subscripts are to be taken mod m. Such a permutation is called an *$\underline{m}$–cycle*.

Notice that the $\underline{m}$–cycle, with $m > 1$, the support of any $\underline{m}$–cycle $(a_1\ a_2 \cdots a_m)$ is the set of letters listed, i.e.

$$\{a_1, \ldots, a_m\}.$$

The exceptional case occurs when $m = 1$ in which case we see that any $\underline{1}$–cycle is just the identity.

Example 7.5. Let's consider the 4 cycle $(1\ 6\ 2\ 4)$ in $\mathcal{S}_7$. In the notation above, it has $a_1 = 1$, $a_2 = 6$, $a_3 = 2$, and $a_4 = 4$. Thus, by definition, it takes

$$\overset{a_1}{\underset{\shortparallel}{1}} \mapsto \overset{a_2}{\underset{\shortparallel}{6}} \mapsto \overset{a_3}{\underset{\shortparallel}{2}} \mapsto \overset{a_4}{\underset{\shortparallel}{4}} \mapsto \overset{a_1}{\underset{\shortparallel}{1}}.$$

Since the letters $3, 5$, and 7 aren't listed, each are fixed. In summary, we have that $(1\ 6\ 2\ 4)$ is the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 6 & 4 & 3 & 1 & 5 & 2 & 7 \end{pmatrix}.$$

- To see where $(a_1\ a_2 \cdots a_m)$ sends a_i, simply look one spot to the right, unless you're already all the way at the end, in which case it goes back around to a_1. This is the reason why its called a cycle.
- Every m–cycle is indeed a permutation (a.k.a. a bijection). Indeed, to find the inverse of an arbitrary m–cycle

  $$(a_1\ a_2 \cdots a_m)$$

  simply read it backwards:

  $$(a_m\ a_{m-1} \cdots a_1).$$

- For an arbitrary m–cycle

  $$\sigma = (a_1\ a_2 \cdots a_m)$$

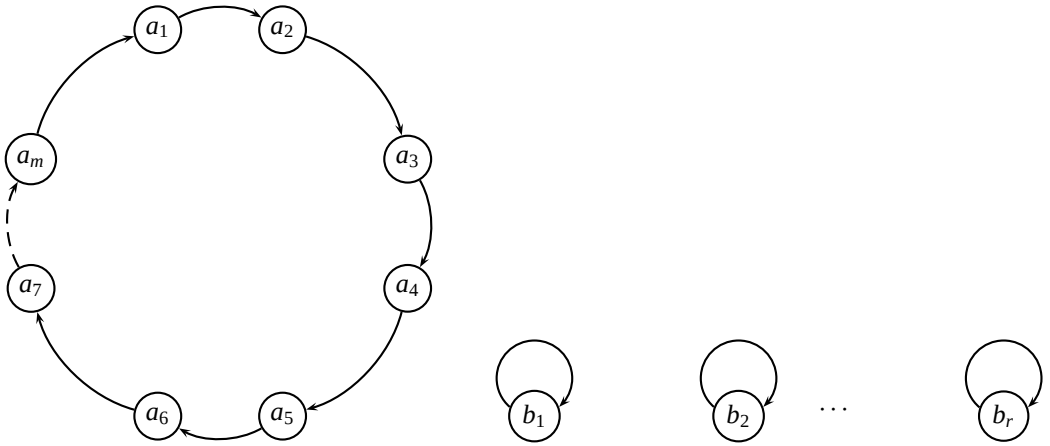
  and k any integer (even negative integers), we have that $\sigma^k[a_i] = a_{i+k}$ (again, subscripts taken mod m). For k positive, this is trivial to prove by induction, the case that $k = 1$ is the definition of an m–cycle while if it were true for $k-1$, then

  $$\sigma^k[a_i] = \sigma\left[\sigma^{k-1}[a_i]\right] = \sigma[a_{i+k-1}] = a_{i+k}.$$

  Using this, the case that k is negative is proven simply by noting that

  $$\sigma^{-k}[a_i] = \cancel{\sigma^{-k}}(\cancel{\sigma^{k}}[a_{i-k}]) = a_{i-k}.$$

- An m–cycle has a cycle decomposition consisting of a single "main" cycle with m vertices and the $n-m$ cycles consisting of singletons. (In the diagram below, to make things look prettier I set $r = n - m$).



Conversely, if σ has a cycle decomposition which looks like this, then σ is indeed the given m–cycle.

- Cycles are independent of the starting point and the order which the trivial cycles are written. For instance the cycles

$$(1\,2\,5\,7\,9\,3) \quad \text{and} \quad (5\,7\,9\,3\,1\,2)$$

are the same, since each is the function (permutation)

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 5 & 1 & 4 & 7 & 6 & 9 & 8 & 3 \end{pmatrix}.$$

More generally, for any integer $i \in \overline{m}$, the $\underline{m}$–cycles

$$(a_1\ a_2 \cdots a_m) = (a_i\ a_{i+1} \cdots a_{m+i})$$

(where all the subscripts are taken mod m) are the same. This is because each denotes the permutation which takes a_k to $a_{k+1 \bmod m}$. The process of changing between the different ways of writing the same cycle is called *recycling*. This matches the fact that the cycle decomposition consists of a single circle (containing more than one vertex) and the vertices of circle really don't have a natural starting point. But since we are dealing with a nicely ordered set, namely the set of integers between 1 and n, we can always rewrite a cycle with its least letter appearing first.

CONVENTION 7.6. By default, we will write all $\underline{m}$–cycles $(a_1 \cdots a_m)$ in *standard form*; this is so that a_1 is the smallest of the a's.

For instance, $(1\,2\,5\,7\,9\,3)$ is in standard form while $(5\,7\,9\,3\,1\,2)$ is not. Also, $(3\,5\,7\,4\,6)$ is in standard form, while $(7\,4\,6\,3\,5\,7)$ is not. Note that we are not requiring that $a_1 < a_2 < a_3$, etc. In fact most permutations cannot be written in that form. For instance, this is the case for the permutation $(1\,3\,2) \in \mathcal{S}_3$. We are only requring that a_1 be the least letter appearing in the cycle.

REMARK 7.7. Again, the analogy is with calling $2/4$ a rational number. This is certainly true, and moreover, this rational number is equal to $1/2$, but we consider the latter to be in standard form. The real advantage is so everybody is speaking the same language!

There are a total of m ways to write the same $\underline{m}$–cycle, but only one way to write it in standard form. The only possibility for confusion comes from the identity $\in \mathcal{S}_n$. Notice that for any letter $i \in \overline{n}$, the identity is the $\underline{1}$–cycle (i). Therefore, we might as well consider *the standard form of the identity* to be the $\underline{1}$–cycle, (1).

EXERCISE 7.1. For $m > 1$, prove that there are a total of

$$\frac{n!}{m \cdot (n-m)!}$$

$\underline{m}$–cycles in $\mathcal{S}_n$. Use this (and perhaps a calculator) to show that in $\mathcal{S}_6$ there are a total of 310 elements which are not cycles.
Hint! We can think of any $\underline{m}$–cycle as m distinct elements of n^{*m}. Use Exercise 2.33 back on page 2.33.

EXERCISE 7.2. If we count the identity, (1) as a $\underline{1}$–cycle, how many elements of $\mathcal{S}_7$ are and are not cycles?[III]

[III] Of course every element of $\mathcal{S}_n$ is a product of (disjoint) cycles, but most (especially for large n) are not actually cycles. For instance $(1\,2)(3\,4)$ is not a cycle, despite the fact that we call it a $\underline{2}\,\underline{2}$–cycle.

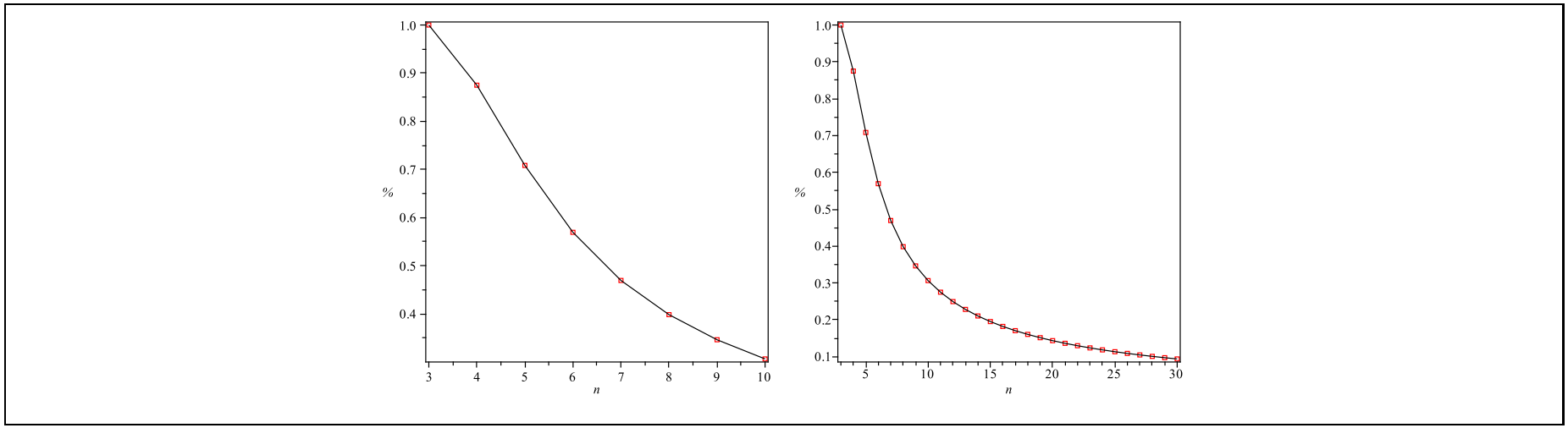


Figure 7.1: The percentage of cycles in $\mathcal{S}_n$. n is plotted along x–axis and $\dfrac{\text{number of cycles in } \mathcal{S}_n}{\text{total number of elements in } \mathcal{S}_n}$ plotted along y–axis. Note that $\mathcal{S}_5$ is the first time there are more non-cycles than cycles.

DEFINITION 7.8. A *non-trivial cycle* is a cycle with more than one letter. We call a permutation itself a cycle if its digraph has only one non-trivial cycle. If that non-trivial cycle contains n letters we call the permutation in question an $\underline{n}$-cycle.

Therefore we technically have two different definitions of a cycle: one definition describes a digraph the other describes a permutation. Notice, however, that one uniquely determines the other. As such we use the same notation to describe both. For instance, the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 1 & 2 \end{pmatrix}$$

is the cycle $(1\ 3\ 5\ 2\ 4)$.

DEFINITION 7.9. Two cycles are *disjoint* iff there supports are disjoint.

Since the support of a cycle is trivial to spot: it's just the set of letters it contains[IV], we it's trivial to spot disjoint cycles, they are simply cycles that have no common letter listed.

EXAMPLE 7.10. The cycles $(1\ 4\ 7\ 9)$ and $(2\ 3\ 5\ 6)$ are disjoint.

[IV] The exception again is a $\underline{1}$–cycle which always has empty support.

Disjoint Circles $\overset{1\text{-}1}{\leftrightsquigarrow}$ Disjoint Cycles.

Figure 7.2: Image for 7.11

THEOREM 7.11. *For a permutation σ, the cycle decomposition of σ consists of mutually disjoint circles $\mathbf{S}_1, \ldots \mathbf{S}_k$, iff $\sigma = \sigma_1 \circ \cdots \circ \sigma_k$ where σ_i are the mutually disjoint cycles corresponding to the circle $\mathbf{S}_i$.*

PROOF. Suppose that the digraph of σ is the disjoint union of circles

$$\mathbf{S}_1, \ldots \mathbf{S}_k$$

with corresponding disjoint cycles

$$\sigma_1, \ldots, \sigma_k.$$

I must show that $\sigma = \sigma_1 \circ \cdots \circ \sigma_k$. To this end let $j \in \overline{n}$ be arbitrary. Now $j \in \overline{n}$ is in one and only one circle, $\mathbf{S}_i$, and so it is moved by only σ_i. Since, by definition of the cycle decomposition of σ, there is an edge from the vertex j to the vertex $\sigma[j]$ we have that $\sigma[j]$ also lives only in this same circle, $\mathbf{S}_i$. Thus, σ_i is also the only cycle listed which moves $\sigma[j]$. As such, we must have that $(\sigma_1 \circ \cdots \circ \sigma_k)[j] = \sigma[j]$. Since these dudes are functions, this means that the functions are equal. ■

COROLLARY 7.12. *Up to the order which they are written, every permutation can be written uniquely as the product (i.e. composition of functions) of (mutually) disjoint cycles.*

PROOF. The digraph of any permutation is the disjoint union of circles. If we fix the order which we draw these circles we fix the order which we write the corresponding permutations. ■

For instance, the permutation discussed in Example 7.3 is

$$(1\ 13\ 5\ 10)(2)(3\ 15\ 8)(4\ 14\ 11\ 7\ 12\ 9)(6).$$

We typically omit trivial cycles from a cycle decomposition. In this case the cycles (2) and (6) are trivial. Therefore, we write the permutation as

$$(1\ 13\ 5\ 10)(3\ 15\ 8)(4\ 14\ 11\ 7\ 12\ 9).$$

For another example, the permutation

$$(1\ 3\ 4\ 6)(2)(5)(7)$$

is a $\underline{4}$-cycle and will typically be written as

$$(1\ 3\ 4\ 6).$$

If we know which degree we are working with, then we can recover the original cycle decomposition simply by putting singeltons (i) at the end of any cycle decomposition. For instance, the permutation

$$(1\ 4) \in \mathcal{S}_4 \qquad \text{really means} \qquad (1\ 4)(2)(3).$$

WARNING! 7.13. Never confuse permutations with letters (i.e. elements of $\overline{n}$)! Permutations are like verbs, they do something. A permutation of a deck of cards is not a noun! On the other hand, the letters, i.e. the arguments of the permutations, are like nouns. While the shuffling of cards is a permutation; you can't "see it" in a still frame, you'd need to see a movie, the actual cards are the letters. If a buddy asked you to shuffle a deck of cards, and you handed them a jack of spades, they'd look at you as if you were crazy. Same thing happens here. For instance, while $2 \in \overline{3}$, but $2 \notin \mathcal{S}_3$.

Definition 7.14. Let σ be a permutation which, when written as a product of disjoint cycles, has the form

$$\sigma_1 \cdots \sigma_k$$

with σ_i a λ_i–cycle. The *cycle type* of σ is the list of numbers $\lambda_1 \ldots \lambda_k$.

Up to adding up additional 1's and the order of the λ_i, there is only one cycle type for a given permutation. We take of the problem with order by making the following convention:

Convention 7.15. The cycle type of a permutation will be written in decreasing order.

The problem of zeros is a bit annoying from a theoretical point of view, but there really will be times when we will want to tack on the 1's and times when we do not wish to tack on the ones. If we include the 1–cycles, we call it the *proper cycle type* of σ. On the other hand, if we do not include any 1's, we call it the *common cycle type*.

Definition 7.16. The set of all permutations in $\mathcal{S}_n$ with (proper) cycle type λ is denoted $\mathrm{cc}(\lambda)$.

Example 7.17. The permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

written in cycle notation is $(1\,2)(3\,4)$. Therefore it has common cycle type 2 2. If we consider it as an element of $\mathcal{S}_4$, than this is also its proper cycle type. However if we consider it as an element of $\mathcal{S}_6$, then its proper cycle type is 2 2 1 1.

There are 3 2 2–cycles in $\mathcal{S}_4$. They are

$$\mathrm{cc}(2\,2) = \{(1\,2)(3\,4), (1\,3)(2\,4), (1\,4)(2\,3)\}.$$

In $\mathcal{S}_5$, however there are 15 2 2–cycles. They are

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

Of course the proper cycle type depends on which n we are in, but the common one does not. Moreover an elements common and proper cycle types are the same if and only if the element does not fix anything. We say that an element *has its cycle type*. For instance m–cycles have (common) cycle type m. A more complicated example is

$(1\ 2\ 3)(4\ 5\ 6)(7\ 8)$ which is a $\underline{3}\,\underline{3}\,\underline{2}$–cycle. If we think of this element as an element of $\mathcal{S}_8$ this is also its proper cycle length. However, if instead we think of this element as an element of $\mathcal{S}_{11}$, then it's proper cycle type is $\underline{3}\,\underline{3}\,\underline{2}\,\underline{1}\,\underline{1}\,\underline{1}$. In this case, we may either say that this element is a $\underline{3}\,\underline{3}\,\underline{2}$–cycle or, like we did before, a $\underline{3}\,\underline{3}\,\underline{2}\,\underline{1}\,\underline{1}\,\underline{1}$–cycle. (So the cycle type is determined up to a tacking on 1's).

WARNING! 7.18. Just because I use the word **cycle** when I call something like $(1\ 2)(3\ 4)$ a $\underline{2}\,\underline{2}$–**cycle**, this element is not cycle at all–it's really the product of disjoint $\underline{2}$–cycles! I use this sloppy language to save time! The only real cycle types are permutations with a only a single number showing up in their cycle type.

CONVENTION 7.19. We think of $\mathcal{S}_n$ as a subset (hence a subgroup) of $\mathcal{S}_{n+1}$ by identifying a permutation of n–letters with a permutation of $n+1$–letters which fixes the last letter.

In symbols, the convention says that $\mathcal{S}_n \preccurlyeq \mathcal{S}_{n+1}$. (This is technically false since elements of the first group are functions on

$$\overline{n} = \{1, \ldots, n\}$$

while elements of the second group are functions on

$$\overline{n+1} = \{1, \ldots, n, n+1\}$$

and there's nothing in common since functions with different domains, are, by definition, not equal to each other.) We should be a little bit careful since letters inside of a single parenthesis such as (3) now have two meanings: it could be part of a cycle in a cycle decomposition which happens to have just a single element, or it could be the argument of some permutations. For instance, $(1\ 2\ 3)(3)$ if it means $(1\ 2\ 3) \circ (3)$ is just the permutation $(1\ 2\ 3)$, while if it means the function $(1\ 2\ 3)$ applied to the letter 3, then the output would be the letter 1. At least initially, we will attempt to limit this potential confusion by writing the latter as $(1\ 2\ 3)[3] = 1$ instead of the now ambiguous $(1\ 2\ 3)(3) = 1$. After this chapter I will assume you can tell from context and this notation will not be used!

I guess this fits what we would think of in our everyday experience. For instance, if I shuffle a deck of cards, I can think of that as a permutation of the cards. Now, if you had taken out the two Jokers from the deck before I shuffled the deck of cards, we could safely say that I also shuffled the entire deck of cards. Nobody would ever bother to tell you "shuffle the entire deck of cards but not the two jokers I spent 5 minutes digging through the deck to find to pull out because if you put them back in the deck and I have to wade through the entire deck to find them again I will go NUTS!" ... ok, I'm getting carried away as usual.

LEMMA 7.20. *An $\underline{m}$-cycle has order m.*

PROOF. Let σ be the $\underline{m}$–cycle $(a_1\ \ldots\ a_m)$. Then, for each a_i in the support of σ we have that

$$\sigma^k[a_i] = a_{(i+k)}.$$

This equals a_i again iff $[\,i+k\,]_m = [\,i\,]_m$ which happens iff $m \mid k$. ∎

EXAMPLE /THEOREM 7.21. Let σ be an m–cycle and k be a positive integer. Define $d = \gcd(m,k)$. Then σ^d is the product of d disjoint $\underline{m/d}$–cycles.

PROOF. For all a_j with $j \in \overline{m}$, define the set

$$\mathbf{O}(a_j) = \{\sigma^{ik}[a_j] : i \in \mathbb{Z}\} = \{a_j, a_{j+d}, \ldots, a_{j+qd}, \ldots, a_{j+(m/d-1)d}\}$$

(all subscripts taken mod m) which contains m/d letters. Notice that since $k = qd$ for some integer q, we have that

$$\sigma^k[a_j] = a_{j+k} = a_{j+qd} \in \mathbf{O}(a_j).$$

Next, we can easily check that

$$\mathbf{O}(a_1), \mathbf{O}(a_2), \ldots, \mathbf{O}(a_d)$$

partition $\{a_1, \ldots, a_m\}$ (why?). Lastly, almost by definition, these d disjoint sets make up the support of the d disjoint cycles whose product is σ^k. ■

For an example of this last phenomena, consider $\sigma = (1\ 4\ 5\ 3\ 8\ 7\ 9\ 2\ 6)$. There we have

$$\sigma^3 = (1\ 3\ 9)(2\ 4\ 8)(5\ 7\ 6).$$

7.2 Getting used to operations in cycle notation

In this section we discuss how to do the basic group theoretical operations in cycle notation.

(1) **(Powers)** Given a permutation σ written in cycle notation, to take the m^{th} power of σ we simply look m spots to the right, where I mean this in the cyclic sense. For instance, for the permutation

$$\sigma = (1\ 4\ 3\ 7)(2\ 5\ 6)$$

we have $\sigma^3 = (1\ 7\ 3\ 4)(2)(5)(6) = (1\ 7\ 3\ 4)$. Actually, for this dude it was easier for the first permutation instead of reading 3 over to simply read one backwards. The reasoning for this is because for an $\underline{m}$ cycle $\sigma = (a_1 \ldots a_m)$, $\sigma^k[a_i] = a_{i+k}$ and a_{i+k} is written in the $\underline{m}$–cycle k spots over (again, I mean this cyclically) from a_i. The same statement works in general since, in general for a group, given two commuting elements a and b, which disjoint cycles are examples of, we have that

$$(a \star b)^k = a^k \star b^k.$$

Note that for the first power, this algorithm states that we can write σ^1 by reading one spot over to the write. However, this is not inconsistent since this is simply recycling the original permutation written in cycle form.

(2) **(Inverses)** Finding the inverse of a permutation written in cycle notation is just as easy, if not easier, than taking powers. Simply read the cycles backwards. For instance, for $\sigma = (1\ 6\ 9\ 8\ 4)(2\ 5\ 3\ 7)$ we have that $\sigma^{-1} = (4\ 8\ 9\ 6\ 1)(7\ 3\ 5\ 2)$ which after recycling is $(1\ 4\ 8\ 9\ 6)(2\ 7\ 3\ 5)$. The reason that this works is because saying that a permutation is the $\underline{m}$–cycle $(a_1 \ldots a_m)$ means that this permutation takes a_i to a_{i+1} (in the cyclic sense). Therefore the inverse takes a_j to a_{j-1}, meaning simply read the permutation backwards. This works in general since the inverse of a product of commuting elements in any group is the product of the inverses.

(3) **(Multiplying)** You would hope that the most basic operation would be to multiply permutations, but this actually turns out to be the hardest! For instance, conjugating a permutation by another turns out to be much, much easier than multiplying them.

Let's consider, by way of example, finding the composition of the three cycles:

$$(1\ 3\ 7\ 6\ 4) \circ (1\ 2\ 5\ 6\ 7\ 4) \circ (1\ 5\ 7\ 4\ 2) \tag{1}$$

Since permutations are functions, we compose from right to left. What does the rightmost permutation do to the letter 1? Well, it takes 1 to 5. The middle permutation takes 5 to 6 (since 5 and 6 are adjacent in this middle cycle) and the leftmost permutation takes 6 to 4. Therefore the composition in (1) has the net effect of taking to 1 to 4. This means if we write this composition as a product of disjoint cycles it starts of with (1 4 . This is illustrated in

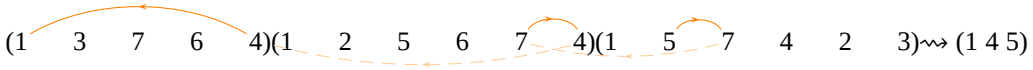


Now to continue the process we now need to figure out what the composition does to 4. Reading things from right to left, we have that 4 goes to 2, 2 goes to 5 and 5 goes to itself. This is illustrated in the second figure.

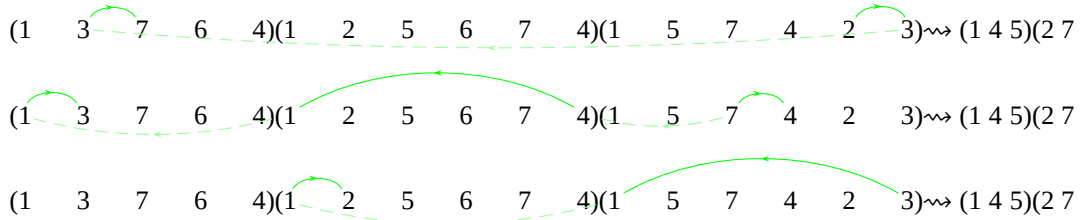


Note that the leftmost permutation, doing nothing to the letter 5, there is no arrow in the first part of the composition. Keeping a running tally we have that 4 goes to 5 and so the permutation which will depict our end result is as of right now (1 4 5 .

Next, we need to investigate the letter 5.[I] Well, we have that 5 goes to 7, 7 goes to 4 and 4 goes back to 1. The net effect therefore takes 5 back to 1. This means, if we were still writing our cycles out as loops, that we have closed the loop. This means that we have completed our first part of the resulting cycle. This is illustrated in the third figure:



What do we do now? Remember, we agreed to start the second cycle with the smallest letter left, which in our case is 2. We do this faster by simply drawing the diagrams:



The only letter that hasn't shown up already (between 1 in 7) is 6. This implies that 6 must go to 6. We illustrate this explicitly because I am just too kind.



EXERCISE 7.3. Let $\sigma = (1\ 2\ 6\ 9)(3\ 5\ 4)$ and $\tau = (1\ 2\ 4)(5\ 9\ 6)$. Perform the following elementary operations (even if you've read the tricks in the next step, please only use the methods from this section):

(1) σ^4 **(3)** τ^{-1} **(5)** $\sigma^{-1}\tau\sigma^2$

(2) τ^2 **(4)** $\sigma\tau\sigma^2$ **(6)** $(\sigma\tau)^4$

SOLUTION

7.3 TRICKS

NOTATION 7.22. For a permutation $\sigma \in \mathcal{S}_n$, the fixed point set of σ is denoted $\overline{n}^{\sigma}$ or $\mathrm{Fix}(\sigma)$ and the support of σ is denoted as $\overline{n}_{\sigma}$ or $\mathrm{Supp}(\sigma)$.

[I] Yes, don't forget they are called **letters**. Cookie monster would not be happy with us at all.

REMARK 7.23. In cycle notation, $\overline{n}^{\sigma}$ are just the set of letters which show up as $\underline{1}$–cycles in σ. Or if we use the reduced notation, these are the dudes who do not show up at all.[I]

EXAMPLE 7.24.

$$\overline{8}^{(1\,8\,5)(2\,4\,6)} = \{3, 7\}$$
$$\overline{5}^{(1\,3)(2\,4)} = \{5\}.$$

In general, we have $\overline{n}^{\sigma} = \overline{n}$ holds if and only if $\sigma = \mathbf{1}_{\overline{n}} = (1)$, the identity of $\mathcal{S}_n$. This is because only the identity permutation fixes all letters. On the other hand, the set of elements which satisfy $\overline{n}_{\sigma} = \overline{n}$, i.e. the set of permutations which fix no letter, is far from trivial. For instance, in the case that $n = 6$ we have

$$\{\sigma : \overline{6}_{\sigma} = \overline{6}\} = \{\underline{6}\text{–cycles}\} \cup \{\underline{3}\,\underline{3}\text{–cycles}\} \cup \underline{4}\,\underline{2}\text{–cycles}\} \cup \{\underline{2}\,\underline{2}\,\underline{2}\text{–cycles}\}.$$

THEOREM 7.25. *For any two permutations $\sigma, \tau \in \mathcal{S}_n$ we have that*

$$\tau\,(\mathrm{Fix}(\sigma)) = \mathrm{Fix}(\tau\sigma\tau^{-1})^{II}$$

and

$$\tau\,(\mathrm{Stab}(\sigma)) = \mathrm{Stab}(\tau\sigma\tau^{-1})$$

PROOF. [III] Since both sets in the second equation are complements of each set of the first equation, we only prove the first equation since we could take complements of both sides of that equation to prove the second equation (or we could prove the second and take complements of both sides to prove the first).

We trying to prove the equality of two sets, $\tau\,(\mathrm{Fix}(\sigma))$ and $\mathrm{Fix}(\tau\sigma\tau^{-1})$. We follow the usual procedure and prove mutual containment.

$\subseteq$ Suppose that

$$i \in \tau(\mathrm{Fix}(\sigma)).$$

Then $i = \tau[j]$ for some $j \in \mathrm{Fix}(\sigma)$. This means that Then $j = \tau^{-1}[i] \in \mathrm{Fix}(\sigma)$. Hence

$$\sigma\left[\tau^{-1}[i]\right] = \sigma[j] = j = \tau^{-1}[i].$$

If we apply τ to both sides of the first and last terms of this equation, we see

$$\tau\left[\sigma\left[\tau^{-1}[i]\right]\right] = (\tau \circ \tau^{-1})[i] = i.$$

That is $i \in \mathrm{Fix}(\tau\sigma\tau^{-1})$.

[II] That is the τ takes the set of letters which are fixed by σ to the set of letters which are fixed by the permutation $\tau\sigma\tau^{-1}$.

[III] This is definitely a left brain proof!

$\supseteq$ All arrows can be turned around to show the reverse inclusion.

■

Example 7.26. Let $\sigma = (1\ 3\ 5)(2\ 6)$. Then $\overline{6}^{\sigma} = \{4\}, \overline{6}_{\sigma} = \{1, 2, 3, 5, 6\}$. Thus if $\tau = (2\ 4\ 6)$, then by the theorem

$$\overline{6}^{\tau\sigma\tau^{-1}} = \tau\{4\} = \{\tau[4]\} = \{6\}$$

and

$$\overline{6}_{\tau\sigma\tau^{-1}} = \tau\{1, 2, 3, 5, 6\} = \{\tau[1], \tau[2], \tau[3], \tau[5], \tau[6]\} = \{1, 4, 3, 5, 2\} = \{1, 2, 3, 4, 5\}.$$

Indeed, we can verify this directly since

$$\tau\sigma\tau^{-1} = (1\ 3\ 5)(2\ 4).$$

Corollary 7.27. *Suppose σ and τ commute. Then for any i fixed by σ we have that $\tau[i]$ is also fixed by σ.*

Proof. Since σ and τ commute, we have that $\tau\sigma\tau^{-1} = \sigma$. By the theorem, we therefore have

$$\tau(\text{Fix}(\sigma)) = \text{Fix}(\tau\sigma\tau^{-1}) = \text{Fix}(\tau\sigma\tau^{-1}).$$

Which is (in symbols) the conclusion of the corollary. ■

Remark 7.28. Quickly, let me give a direct proof of this. This might be useful since many students have a tough time with the more general statement just given. Suppose that i if fixed by σ. For emphasis, let's denote by j the image of i under τ, i.e. $j := \tau[i]$. Then we have that

$$\begin{aligned} \sigma[j] &= \sigma[\tau[i]] \\ &= \tau[\sigma[i]] && (\tau \text{ and } \sigma \text{ commute.}) \\ &= \tau[i] && (i \in \overline{n}^{\sigma}) \\ &= j && j := \tau[i]. \end{aligned}$$

Thus, by definition, j if fixed by σ.

For instance we can tell simply by looking that the permutations

$$(1\ 2\ 3\ 4\ 5) \quad \text{and} \quad (1\ 4\ 6)(2\ 3\ 5)$$

do not commute. This is because the first permutation fixes the letter 6, the second takes 6 to 1, but the first permutation does not fix 1. Of course the converse if false. For instance the permutation $(1\ 2\ 3\ 5\ 4)$ also does not commute with $(1\ 2\ 3\ 4\ 5)$ even though both leave each other's fixed point, 6, alone.

Theorem 7.29. *Suppose $\Omega \leqslant \mathcal{S}_n$. Assume for every $i \in \overline{n}$ there exists a permutation $\zeta_i \in \Omega$ with the property that ζ_i fixes only the letter i. Then Ω is centerless.*

Proof. Suppose that $\sigma \in \Omega$ is central. We need to prove that $\sigma = (1)$. Since the elements of $\mathcal{S}_n$ are functions, this is the same thing as proving that σ fixes every element of $\overline{n}$ (its domain); i.e. proving $\forall i \in \overline{n}, \sigma[i] = i$. Let $i \in \overline{n}$ be an arbitrary letter. Then, by assumption, there exists an element $\zeta_i \in \Omega$ who only fixes the letter i. Since σ is assumed to be central in Ω, and ζ_i is an element of Ω, σ and ζ_i commute. Therefore, Corollary 7.27 applies to show that

$$\sigma[\{i\}] = \sigma[\text{Fix}(\zeta_i)] = [\text{Fix}(\zeta_i)] = \{i\}.$$

This means that $\sigma[i] = i$. Since $i \in \overline{n}$ was arbitrary, $\sigma = (1)$ which is what we wanted to prove. ■

Exercise 7.4. Prove that for $n > 2$, $\mathcal{S}_n$ is centerless.
Hint! Consider permutations, which in the special case $n = 4$ are $\zeta_1 := (2\ 3\ 4)$, $\zeta_2 := (1\ 3\ 4)$, $\zeta_3 := (1\ 2\ 4)$, and $\zeta_4 := (1\ 2\ 3)$. Apply the theorem, Theorem 7.29.

SOLUTION

In another example of the use of the theorem, if we think of $\mathcal{D}_5$ as a subset of $\mathcal{S}_5$ (after labeling the vertices of a regular pentagon and thinking of an element of $\mathcal{D}_5$ as permuting the vertices) the 5 reflections each fix exactly 1 vertex. Thus $\mathcal{S}_5$ is centerless.

Exercise 7.5. Generalize this comment to prove that $\mathcal{D}_n$ is centerless whenever n is odd.

SOLUTION

The following definition is the obvious generalization of Definition 7.9.

Definition 7.30. Two permutations are said to be *disjoint* if their supports do not intersect.

Written out this means that σ and τ are disjoint if

$$\sigma[i] \neq i \Rightarrow \tau[i] = i.$$

In terms of reduced cycle decompositions, it's trivial to spot disjoint permutations; they are written with no letters in common. This is because the support of a permutation consists precisely of the set of letters which show up in a (reduced) cycle decomposition.

Example 7.31.

(1) The identity, $\mathbf{1}_{\overline{n}} = (1)$, having empty support, is disjoint from all permutations, even itself.[IV]

(2) The permutations $(1\ 5\ 8)(3\ 7)$ and $(2\ 4)(6\ 9)$ are disjoint since

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

(3) The permutations $(1\ 6\ 5\ 3\ 8)$ and $(2\ 7)(4\ 9\ 5)$ are not disjoint since their supports have the letter 2 in common.

(4) More generally, if the support of a permutation $\sigma \in \mathcal{S}_n$ contains m letters, then it is disjoint from exactly $(n - m)!$ other permutations.

[IV] Clearly this is the only permutation with this property since it is the only permutation with empty support.

(5) It's easy to see that two permutations are disjoint from each other iff all any cycle in the cycle decomposition of the first is disjoint from any cycle in the cycle decomposition of the second.

The following theorem should be intuitive. For instance, if we take a deck of cards and cut that deck in half it doesn't matter which half we shuffle first–the net shuffling is the same.

THEOREM 7.32. *Disjoint permutations commute.*

PROOF. Suppose σ and τ are disjoint cycles. By DeMorgan's law,

$$\overline{n}_\sigma \cap \overline{n}_\tau = \varnothing \Rightarrow \overline{n}^\sigma \cup \overline{n}^\tau = \overline{n}. \tag{2}$$

In words, every letter is fixed by either σ or τ. To show that $\sigma\tau = \tau\sigma$, we need to show that both sides do the same thing to every letter. Let $i \in \overline{n}$. There are three cases:

Case 1 Suppose that

$$i \in \overline{n}^\sigma \cap \overline{n}^\tau.$$

That is we are assuming that both σ and τ fix the letter i. Therefore we have

$$\sigma[\tau[i]] = \sigma[i] = i \quad \text{and} \quad \tau[\sigma[i]] = \tau[i] = i.$$

Case 2 Suppose that

$$i \in \overline{n}^\sigma \cup \overline{n}_\tau.$$

That is we are assuming that σ fixes i but τ moves it to say $j \neq i$. Then I claim that τ also moves j. Indeed, since τ is one-to-one and since $j \neq i$ we must have that $\tau[j] \neq \tau[i] = j$. So by assumption we have that σ fixes j (in addition to i). Hence

$$\tau[\sigma[i]] = \tau[i] = j \qquad \text{while} \qquad \sigma[\tau[i]] = \sigma[j] = j.$$

Thus we have that $(\tau\sigma)[i] = (\sigma\tau)[i]$ in this case.

Case 3 The final case that

$$i \in \overline{n}_\sigma \cup \overline{n}^\tau.$$

is logically the same as the previous case.

Hence in all cases, we have that $(\tau\sigma)[i] = (\sigma\tau)[i]$. Since $\tau\sigma$ and $\sigma\tau$ are two permutations (which are functions) which do the same thing to every element of their domain, we have that these two permutations (which are functions) are equal. Therefore σ and τ commute. ■

REMARK 7.33. The converse–that only disjoint permutations commute–is obviously false. For instance, any non-identity permutation is not disjoint from itself, but any element of any group commutes with itself. Less trivially, we have that $(1\,2)(3\,4)$ and $(1\,3)(2\,4)$ commute but are not disjoint.

EXERCISE 7.6. Let σ be an $\underline{n}$–cycle in either $\mathcal{S}_n$ or $\mathcal{S}_{n+1}$.

SOLUTION

(1) Prove that if τ commutes with σ, then τ is a power of σ, i.e. $\tau = \sigma^m$.[V]

(2) Prove for $m > n + 1$ we always have that

$$C_{\mathcal{S}_m}(\sigma) \neq \langle \sigma \rangle .$$

Hint! In words, I'm asking you to prove that in this case there exists a permutation $\tau \in \mathcal{S}_m$ which commutes with σ but which is not a power of σ.

So, in the cases of the above exercise, only powers of an n–cycle commute with it. More generally:

EXERCISE 7.7. Let σ be an m–cycle in $\mathcal{S}_n$ ($m \leq n$).

SOLUTION

(1) Prove that every permutation which commutes with σ has the form

$$\sigma^i \cdot \tau \tag{3}$$

where i is an integer and τ is a permutation disjoint from σ.[VI]

(2) Prove that the centralizer of σ contains $m \cdot (n - m)!$ elements.

THEOREM 7.34. *If σ is a $\lambda_1 \ldots \lambda_m$–cycle, then*

$$|\sigma| = \operatorname{lcm}(\lambda_1, \ldots, \lambda_m).$$

PROOF. We hope to be able to apply Proposition 5.58 back on page 279. To do this we need to very two things.

Suppose that σ is written as a product of two disjoint permutations $\sigma_1\sigma_2$. Since

$$\overline{n}_{\sigma_1} \cap \overline{n}_{\sigma_2} = \varnothing$$

by assumption, we have that $\langle \sigma_1 \rangle \cap \langle \sigma_2 \rangle$ is trivial. This is because every power of a permutation has support contained inside of the original permutation. Moreover we have already shown that this hypothesis also implies that they commute. Thus, by the Lemma, Lemma 5.58, we have that

$$|\sigma| = |\sigma_1\sigma_2| = \operatorname{lcm}(|\sigma_1|, |\sigma_2|).$$

The proof now can now be completed by induction while using the fact that an m–cycle has order m, and the fact that for integers $k_1, \ldots, k_m$, we have that

$$\operatorname{lcm}(\operatorname{lcm}(k_1, \ldots, k_{m-1}), k_m) = \operatorname{lcm}(k_1, \ldots, k_m).$$

■

EXAMPLE 7.35.

(1) $|(1\ 3\ 5)(2\ 4)| = \operatorname{lcm}(3, 2) = 6$

[V] Symbolically, I'm asking you to prove

$$C_{\mathcal{S}_n}(\sigma) = \langle \sigma \rangle = C_{\mathcal{S}_{n+1}}(\sigma).$$

(2) $|(1\ 4\ 6\ 8\ 5\ 2)(3\ 7\ 9\ 10)| = \operatorname{lcm}(6,4) = 12$
(3) $|(1\ 2)(3\ 4\ 5)(6\ 7\ 8\ 9\ 10)| = \operatorname{lcm}(2,3,5) = 30$
(4) $|(1\ 2\ 3)(4\ 5\ 6\ 7\ 8)(9\ 10\ 11\ 12\ 13\ 14\ 15)| = \operatorname{lcm}(3,5,7) = 105.$

EXERCISE 7.8. Find elements of $\mathcal{S}_{16}$ or order 63, 66, 120, and 140.

SOLUTION

EXERCISE 7.9. Prove that for $n < 16$, that $\mathcal{S}_n$ does not have an element of order 63, 66, 120, or 140. (It's fine if you hand–wave a bit with the number theory.)

SOLUTION

EXERCISE 7.10. Find the order of

$$([2]_{15}, [18]_{15}, (1\ 2\ 3\ 4)(5\ 6\ 7)) \in \mathcal{U}(\mathbb{Z}_{15}) \times \mathbb{Z}_{30} \times \mathcal{S}_7.$$

SOLUTION

EXERCISE 7.11. What is the order of an element of largest order in $\mathcal{S}_4 \times \mathcal{S}_4$?

SOLUTION

DEFINITION 7.36. A permutation σ is *homogenous* if it's common cycle type is constant.

For example, since there's only a single number in the cycle type of a cycle, cycles are homogenous. For another example, we can see directly that every element of $\mathcal{S}_4$ is homogenous. However, in $\mathcal{S}_5$ the $\underline{3}\,\underline{2}$-cycle $(1\ 2\ 3)(4\ 5)$ is not homogenous (simply because $3 \neq 2$).

EXERCISE 7.12. Prove that a non-identity permutation σ is homogenous iff for every power of σ^N, $N \in \mathbb{Z}$ which does not equal the identity, we have that

$$\operatorname{Supp}(\sigma) = \operatorname{Supp}(\sigma^N).$$

SOLUTION

EXERCISE 7.13. Prove that if σ is homogenous, then so is every power of σ.

SOLUTION

EXAMPLE 7.37.

(1) The element $(1\ 2)$ and $(2\ 3)$ are conjugate in $\mathcal{S}_3$ since $(1\ 2\ 3)(1\ 2)(1\ 2\ 3)^{-1} = (2\ 3)$.

(2) The elements $(1\ 2\ 3)$ and $(1\ 2)$ are not conjugate since

$$\begin{aligned} (1)(1\ 2)(1)^{-1} &= (1\ 2) & (1\ 2)(1\ 2)(1\ 2)^{-1} &= (1\ 2) \\ (2\ 3)(1\ 2)(2\ 3)^{-1} &= (1\ 3) & (1\ 2\ 3)(1\ 2)(1\ 2\ 3)^{-1} &= (2\ 3) \\ (1\ 3)(1\ 2)(1\ 3)^{-1} &= (2\ 3) & (1\ 3\ 2)(1\ 2)(1\ 3\ 2)^{-1} &= (1\ 3). \end{aligned}$$

As you should guess, there is certainly much better ways of doing this besides testing every element like we just did! We will get to them shortly.

(3) In an abelian group, an element is conjugate to itself and nothing else.

THEOREM 7.38.

(1) **(Cut and Paste)** *Suppose that* $\sigma = (a_1 \cdots a_n)$ *and* $\tau = (a_n \cdots a_{n+m})$ *have* (*supports*) *only* a_n *in common. Then*

$$\sigma\tau = (a_1 \cdots a_{n+m})$$

(2) **($\mathcal{S}_n$ Conjugation Trick)** *To get the cycle decomposition of* $\sigma\tau\sigma^{-1}$ *simply replace every letter* i *in the cycle decomposition of* τ *with* $\sigma[i]$.

PROOF.

(1) For $n \leq i < n+m$ we have that $\tau[a_i] = a_{i+1}$. Since this implies that $n < i+1 \leq n+m$ we have that a_{i+1} does not live in the support of σ and $\sigma[\tau[a_i]] = a_{i+1}$. We get the same conclusion if we take $1 \leq i < n$ (this time the a_i will not be in the support of τ). The only case left is the case that $i = n+m$. But in this case τ takes a_i to a_n and then σ takes a_n back to a_1.

(2) {Draw two parallel spokes } There is an arrow from the letter i to the letter j in the cycle decomposition of τ iff $\tau[i] = j$. Now, if this is the case, then

$$\sigma\tau\sigma^{-1}[\sigma[i]] = \sigma\tau[i] = \sigma[j].$$

Hence, there is an arrow from the letter $\sigma[i]$ to the letter $\sigma[j]$ in the cycle decomposition of $\sigma\tau\sigma^{-1}$. ■

In words, the $\mathcal{S}_n$ conjugacy trick says that **when we conjugate we replace**. It's also cool because it says that, rather surprisingly, that conjugating in $\mathcal{S}_n$ is easier than multiplying! The cut and paste business becomes more believable with the more problems one does. Try a few examples! For instance, directly multiply out (i.e. compose)

$$(1\ 2\ 4\ 6\ 5)(3\ 5\ 7) = (1\ 2\ 4\ 6\ 5)(5\ 7\ 3)$$

and see that you get $(1\ 2\ 4\ 6\ 5\ 7\ 3)$. However, the other part, in my opinion, doesn't get any more intuitive with worked problems!

EXAMPLE 7.39. If we let $\sigma = (1\ 4\ 3\ 7\ 6)$ and $\tau = (1\ 2\ 6)(3\ 5)$, then we have

$$\sigma\tau\sigma^{-1} = (1\ 4\ 3\ 7\ 6)\,[(1\ 2\ 6)(3\ 5)]\,(1\ 4\ 3\ 7\ 6)^{-1} = (\sigma[1]\ \sigma[2]\ \sigma[6])(\sigma[3]\ \sigma[5]) = (4\ 2\ 1)(7\ 5)$$

which after recycling is equal $(1\ 4\ 2)(5\ 7)$. If you want to see a picture

$$(1\ 4\ 3\ 7\ 6) \qquad \begin{array}{ccccc} (1 & 2 & 6) & (3 & 5) \\ | & | & | & | & | \\ \sigma[1]=4 & \sigma[2]=2 & \sigma[6]=1 & \sigma[3]=7 & \sigma[5]=5 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ (4 & 2 & 1) & (7 & 5) \end{array} \qquad (1\ 4\ 3\ 7\ 6)^{-1}$$

EXERCISE 7.14. Suppose that $a_1, \ldots, a_m$ are distinct letters in $\overline{n}$. Show that

$$(a_1\ a_2)\cdots(a_{m-1}\ a_m) = (a_1 \cdots a_m).$$

Show more generally that if $1 < n_1 < n_2 < \ldots < n_k < m$, then

$$(a_1\ a_2 \cdots a_{n_1-1}\ a_{n_1})(a_{n_1}\ a_{n_1+1} \cdots a_{n_2-1}\ a_{n_2}) \cdots (a_{n_k}\ a_{n_k+1} \cdots a_{m-1}\ a_m) = (a_1 \cdots a_m).$$

Hint! By using associativity it is easy to break the second case down into the first case.

SOLUTION

EXERCISE 7.15. Use the previous exercise to compute the following products (i.e. write the products as products of disjoint cycles)

$$(1\ 4\ 2\ 6)(6\ 7\ 8\ 3\ 9) \qquad (1\ 4\ 2\ 6\ 9)(3\ 6\ 7\ 8) \qquad (1\ 2\ 4)(2\ 5\ 9)(5\ 7\ 8).$$

Verify your results by multiplying these out the longer way (i.e. chasing around the letters 1 through 9).

SOLUTION

EXERCISE 7.16. Show that

$$(1\ 2)(1\ 3)\cdots(1\ m) = ((1\ 2 \cdots m))^{-1} = (m \cdots 2\ 1).$$

SOLUTION

THEOREM 7.40. *Two elements of $\mathcal{S}_n$ are conjugate iff they have the same cycle type.*

PROOF.

$\Rightarrow$ The $\mathcal{S}_n$ conjugation trick says that conjugation is the same as replacement. Hence conjugation cannot change the cycle type of a permutation.

⇐ For ease of notation, we illustrate this only for a $\underline{3}\,\underline{2}\,\underline{1}$–cycle in $\mathcal{S}_6$. If you understand this proof, the general proof should be clear. Let τ be an arbitrary permutation of this type. Then τ can be written as the product of disjoint cycles:

$$\tau = (a\ b\ c)(d\ e)(f).$$

Since $\sim_{\mathcal{S}_6}$ is an equivalence relation, we only need to show that τ is conjugate to the standard permutation

$$(1\ 2\ 3)(4\ 5)(6)$$

of this cycle type to be able to conclude that all permutations of this cycle type are conjugate to each other.

We now define σ to be the permutation (written in the original notation from Section 2.3.1 back on page 36)

$$\sigma := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ a & b & c & d & e & f \end{pmatrix}.$$

Then by the $\mathcal{S}_n$ conjugation trick,

$$\sigma \circ (1\ 2\ 3)(4\ 5)(6) \circ \sigma^{-1} = (\sigma[1]\ \sigma[2]\ \sigma[3])(\sigma[4]\ \sigma[5])(\sigma[6]) = (a\ b\ c)(d\ e)(f). \qquad \blacksquare$$

For instance, this says that any two cycles of the same length are conjugate.

EXAMPLE 7.41. To illustrate the theorem, notice that the $\underline{3}$–cycle $(2\ 5\ 7)$ is conjugate to $(1\ 2\ 3)$ since (using the permutation $(1\ 2)(2\ 5)(3\ 7)$))

$$[(1\ 2)(2\ 5)(3\ 7)]\ (1\ 2\ 3)\ [(1\ 2)(2\ 5)(3\ 7)]^{-1} = (2\ 5\ 7).$$

In general, it's quite easy to figure out what permutation to use to do the conjugation. Simply line up the permutations! For instance, consider the permutations $\tau_1 := (1\ 6\ 4\ 8\ 3)(2\ 5\ 7)$ and $\tau_2 := (3\ 7\ 8\ 4\ 5)(1\ 2\ 6)$.

$$\frac{(1\ 6\ 4\ 8\ 3)(2\ 5\ 7)}{(3\ 7\ 8\ 4\ 5)(1\ 2\ 6)} \qquad \rightsquigarrow \qquad \sigma := \begin{pmatrix} 1 & 6 & 4 & 8 & 3 & 2 & 5 & 7 \\ 3 & 7 & 8 & 4 & 5 & 1 & 2 & 6 \end{pmatrix} = (1\ 3\ 5\ 2)(4\ 8)(6\ 7).$$

Then taking this σ we have $\sigma \circ \tau_1 \circ \sigma^{-1} = \tau_2$.

EXERCISE 7.17. For any two $\sigma, \tau \in \mathcal{S}_n$ show that $\sigma\tau$ has the same cycle type as $\tau\sigma$ does.

SOLUTION

7.4 CYCLE TYPES AND PARTITIONS OF AN INTEGER

DEFINITION 7.42. Let n be a positive integer. A partition of n is a listing, λ, of positive integers which add up to n. For instance the (proper) cycle type of a permutation in $\mathcal{S}_n$ is a partition of n.

One typically writes $\boldsymbol{\lambda}$ in increasing (or decreasing if you prefer) order

$$\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m) \qquad \lambda_1 \geq \cdots \geq \lambda_m.$$

For instance, since $5 + 3 + 1 + 1 = 10$, $(5, 3, 1^2)$ is a partition of 10. Other partitions of 10 include (2^5) and $(4, 3, 2, 1)$. While

$$\begin{array}{ll} (1^5) \leftrightsquigarrow 1 \geq 1 \geq 1 \geq 1 \geq 1 & (2, 1^3) \leftrightsquigarrow 2 \geq 1 \geq 1 \geq 1 \\ (3, 1^2) \leftrightsquigarrow 3 \geq 1 \geq 1 & (4, 1) \leftrightsquigarrow 4 \geq 1 \\ (5) \leftrightsquigarrow 5 & (2^2, 1) \leftrightsquigarrow 2 \geq 2 \geq 1 \\ (3, 2) \leftrightsquigarrow 3 \geq 2 & \end{array} \tag{4}$$

make up all 7 partitions of 5.[I]

NOTATION 7.43. The number of partitions of an integer n is denoted $\pi(n)$.

Since the numbers are fairly small, we can see directly that $\pi(1) = 1$, $\pi(2) = 2$, $\pi(3) = 3$, $\pi(4) = 5$, $\pi(5) = 7$, and $\pi(6) = 11$.

As indicated, to save space, we often use the notations on the left hand sides of (4). When this is done, as in

$$\boldsymbol{\lambda} = (\lambda_1^{n_1}, \lambda_2^{n_2}, \ldots, \lambda_m^{n_m})$$

we insist that the inequalities, $\lambda_i \geq \lambda_{i+1}$, are strict. For instance,

$$(5, 3^3, 2^2, 1^3) \leftrightsquigarrow 5 \geq 3 \geq 3 \geq 3 \geq 2 \geq 2 \geq 1 \geq 1 \geq 1$$

of $5 + 3 \cdot 3 + 2 \cdot 2 + 3 \cdot 1 = 21$.

Partitions of n index many things in math. In this course alone we'll see two things that they index. The first we have already seen, and that's proper cycle types, and the other thing is the number of distinct abelian groups of order p^n, c.f. Corollary 13.51 coming up on page 709.

$$\{\text{Proper cycle types in } \mathcal{S}_n\} \leftrightsquigarrow \{\text{ Partitions of } n\} \leftrightsquigarrow \{\text{Conjugacy Classes in } \mathcal{S}_n\}$$

This is all due to the fact that for any $\boldsymbol{\lambda}$–cycle σ, we have

$$\mathfrak{cc}_{\mathcal{S}_n}(\sigma) = \mathfrak{cc}(\boldsymbol{\lambda}).$$

While there isn't a formula per se, there is a fairly straight forward way (especially with a calculator) of calculating $\pi(n)$. We now outline this and pump out the number of partitions of all numbers 1 through 52.

Let

$$q(x) = \sum_{i=0}^{\infty} x^i.$$

From basic calculus we know that q is a geometric series and hence equal to

$$\frac{1}{1-x}.$$[II]

Indeed this is easy to see directly, just multiply both sides by $(1 - x)$. Then we have (by expanding the power series)

[I] For now, this is proof by "I can't think of anymore!"

[II] On its radius of convergence.

THEOREM 7.44.

$$\prod_{i=1}^{\infty} \frac{1}{1-x^i} = \pi(0) + \pi(1)\,x + \pi(2)\,x^2 + \cdots = \sum_{n=0}^{\infty} \pi(n)\,x^n.$$

This theorem is much more useful than it seems since $q(x^j)$ can be truncated as follows:

$$\frac{1}{1-x^j} = 1 + x^j + x^{2j} + x^{3j} + \cdots .$$

And as soon as we have an integer m so that $mj > n$, it cannot possibly contribute to the coefficient of x^n. So we have that

DEFINITION 7.45. For a real number $r > 0$, define the *floor* of r, denoted $\lfloor r \rfloor$, is defined to be the largest integer n so that $n \le r$ and the *ceiling* of r, denoted $\lceil r \rceil$, is defined to be the smallest integer n so that $r \le n$.

Note for every real number r, $\lfloor r \rfloor + \epsilon = \lceil r \rceil$ where $\epsilon = 1$ if r is not an integer, and $\epsilon = 0$ if on the other hand r is an integer.

COROLLARY 7.46. *For every positive integer n*

$$\sum_{i=1}^{n} \pi(i)\,x^i = \prod_{i=1}^{n} \sum_{j=1}^{\lfloor \frac{n}{i} \rfloor} x^j.$$

The advantage of this is that this is something that we can handle by hand and avoid proofs by "I can't think of anymore."

The proofs of these theorems are obvious to anybody who has a background in combinatorics. If you don't, it's best understood by experimenting a bit and giving some thought to what it means to multiply polynomials (of course you can truncate the polynomials so they have no powers greater than n). For instance, in the polynomial

$$(1 + x + x^2 + x^3 + x^4)(1 + x^2 + x^4)(1 + x^3)(1 + x^4)$$

what is the coefficient of x^4? Well, if we take x^4 from the first term in the product, then we'd have to take 1's in the other 3 terms. This is reflected in the fact that $1 + 1 + 1 + 1 = 4$. How else can we get x^4? Well, if we take x^3 from the first term, we're screwed because there is no way to get us up to x^4. How about if instead we took x^2? Then in the second term we'd have to take the x^2 term from there. This is reflected in the fact that $2 + 1 + 1 = 4$. Moving on if we took x from the first we would have to take x^3 from the third. If we took 1 from the first, we could take either x^4 from the second (reflecting the fact that $2 + 2 = 4$) or the x^4 from the last, which reflects the fact that $4 = 4$.

We will only use this result to get some intuition about things and its proof is only of separate interest.

For instance, entering the command

```
> series(product(1/(1-x^i),i=1..53),x,53);
```

into `MAPLE` produces

$$\begin{aligned} 1 + x + 2\,x^2 + 3\,x^3 + 5\,x^4 + 7\,x^5 + 11\,x^6 + 15\,x^7 + 22\,x^8 + 30\,x^9 + 42\,x^{10} + 56\,x^{11} + 77\,x^{12} \\ + 101\,x^{13} + 135\,x^{14} + 176\,x^{15} + 231\,x^{16} + 297\,x^{17} + 385\,x^{18} + 490\,x^{19} + 627\,x^{20} \\ + 792\,x^{21} + 1002\,x^{22} + 1255\,x^{23} + 1575\,x^{24} + 1958\,x^{25} + 2436\,x^{26} + 3010\,x^{27} \\ + 3718\,x^{28} + 4565\,x^{29} + 5604\,x^{30} + 6842\,x^{31} + 8349\,x^{32} + 10143\,x^{33} + 12310\,x^{34} \\ + 14883\,x^{35} + 17977\,x^{36} + 21637\,x^{37} + 26015\,x^{38} + 31185\,x^{39} + 37338\,x^{40} \\ + 44583\,x^{41} + 53174\,x^{42} + 63261\,x^{43} + 75175\,x^{44} + 89134\,x^{45} + 105558\,x^{46} \\ + 124754\,x^{47} + 147273\,x^{48} + 173525\,x^{49} + 204226\,x^{50} + 239943\,x^{51} + 281589\,x^{52} + O(x^{53}) \end{aligned} \quad (5)$$

(the last term $O(x^{53})$ should be thought of as "dot dot dot") which shows (for instance) that $\pi(34) = 12,310$, $\pi(45) = 89134$, and $\pi(52) = 281,589$, etc. This means, for instance, with cards facing down, there are 281,589 ways one can divide a deck of cards up into piles. Alternatively, this means there are 281,589 conjugacy classes in $\mathcal{S}_{52}$.

For low numbers, these numbers assure that after finding a certain number of partitions of a natural number n, we have found them all; as such it saves us from the dreaded .. ummmm, "I can't find anymore" kinds of proofs that we used in the case that $n = 5$. We end the section with a table which makes the information in the above polynomial seem a bit more organized. (To find this table faster we rerecorded it on page 881 in the appendix.)

n	1	2	3	4	5	6	7	8	9	10
$\pi(n)$	1	2	3	5	7	11	15	22	30	42
n	11	12	13	14	15	16	17	18	19	20
$\pi(n)$	56	77	101	135	176	231	297	385	490	627
n	21	22	23	24	25	26	27	28	29	30
$\pi(n)$	792	1002	1255	1575	1958	2436	3010	2718	4565	5604
n	31	32	33	34	35	36	37	38	39	40
$\pi(n)$	6842	8349	10143	12310	14883	17977	21637	26015	31185	37338
n	41	42	43	44	45	46	47	48	49	50
$\pi(n)$	44583	53174	63261	75175	89134	105558	124754	147273	173525	204226

Table 7.1: Partitions of n

Fo r instance $\pi(21) = 792$ gives us some idea of how many ways there are of getting 21 while playing blackjack. (This isn't strictly accurate since an ace can count as 1 or 11 in that game, we only have 4 of every card of face value less than 10, as well as a few other issues.)

EXERCISE **7.18.** As usual for new concepts, you should at least glance at wiki for some history/context of this function. The relevant link is

http://en.wikipedia.org/wiki/Partition (number theory)

SOLUTION

7.5 LARGEST POSSIBLE ORDER AN ELEMENT OF $\mathcal{S}_n$ CAN HAVE*

It's interesting to ask ourselves what the largest order that is possible for an element of $\mathcal{S}_n$ for a given n. We denote this number as $g(n)$. In the literature, it is known as the *Landau function*. For a partition of n,

$$\boldsymbol{\lambda} = \lambda_1 \geq \cdots \geq \lambda_k$$

define $\text{lcm}(\boldsymbol{\lambda})$ as the least common multiple of the entries, λ_i. That is

$$\text{lcm}(\boldsymbol{\lambda}) = \text{lcm}(\lambda_1, \ldots, \lambda_k).$$

We are now able to rephrase this question as follows:

PROPOSITION 7.47.

$$g(n) = \max\{\text{lcm}(\lambda) : \lambda \text{ is a partition of } n\}.$$ ■

This should feel almost like a Calculus 3 problem of finding a maximum of a function, namely the least common multiple, given a constraint, namely that the sequence adds up to n. However this is far most difficult since we can not use calculus to help us solve our current problem!

RANT 7.48. This intuition comes from trial and error. It seems as if though the more "spread out" you can take your entries, the larger the least common multiple will be. Moreover, it seems fairly clear that your best bet is to take the elements of the partition as primes, or at least powers of primes. Formally, this is because for distinct primes, $p_1 \dots p_k$,

$$\text{lcm}(p_1^{k_1}, p_2^{k_2}, p_3^{k_3}, \dots, p_m^{k_m}) = \text{lcm}(p_1^{k_1} p_2^{k_2}, \dots, p_m^{k_m})$$

since both sides are equal to the product $p_1^{k_1} \cdots p_m^{k_m}$, but in terms of sums

$$p_1^{k_1} + p_2^{k_2} + p_3^{k_3} + \cdots + p_m^{k_m} < \left(p_1^{k_1} \cdot p_2^{k_2}\right) + p_3^{k_3} + \cdots + p_m^{k_m}.$$

For instance, if you have a 6 in your partition, which of course adds 6 to your sum, you would have been better off adding in a 2 and 3 to your partition, since there you still have the same effect on least common multiples, but your sum has one less term (since $2 + 3 = 5 < 6$) and you have one extra free number to add to another term.

Meanwhile, you certainly shouldn't have any repeated primes in your partition. This is because

$$\text{lcm}(p, p, p_1, \dots, p_k) = \text{lcm}(p, p_1, \dots, p_k)$$

and deleting the extra p takes away p from the sum, which you certainly can use for something else.

MAPLE EXAMPLE 7.49. The MAPLE code

```
>  with(combinat):
>  PP:=proc(n)
    local i, flag:
    flag:={}:
    for i in partition(n,n) do flag:=flag union {lcm(op(convert(i,set)))} od:
    return(flag):
  >end:
  >G:=n->max(PP(n));
```

Gives us a command, G, when fed an integer n spits out $g(n)$. For instance,

```
>G(10);
```

returns 30. This seems reasonable since, for one, we can find an element, namely $(1\,2\,3\,4\,5)(6\,7\,8)(9\,10)$ in $\mathcal{S}_{10}$ whose order is 30, and it doesn't seem possible to find a partition of 10 with a larger least common multiple.

Producing a sequence, via

```
>seq(G(n),n=1..30)
```

spits out (after a considerable amount of time!) the following list

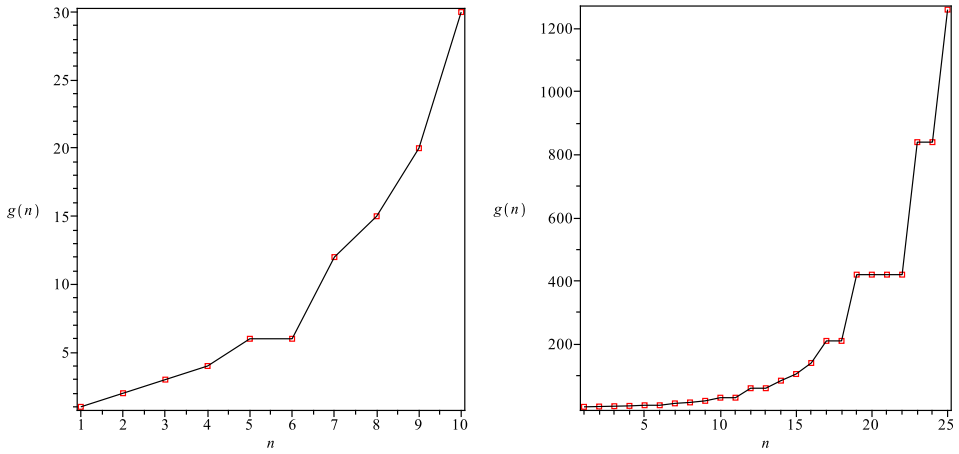
$$1, 2, 3, 4, 6, 6, 12, 15, 20, 30, 30, 60, 60, 84, 105, 140, 210, 210, 420,$$
$$420, 420, 420, 840, 840, 1260, 1260, 1540, 2310, 2520, 4620.$$

Making this into an awesome table:

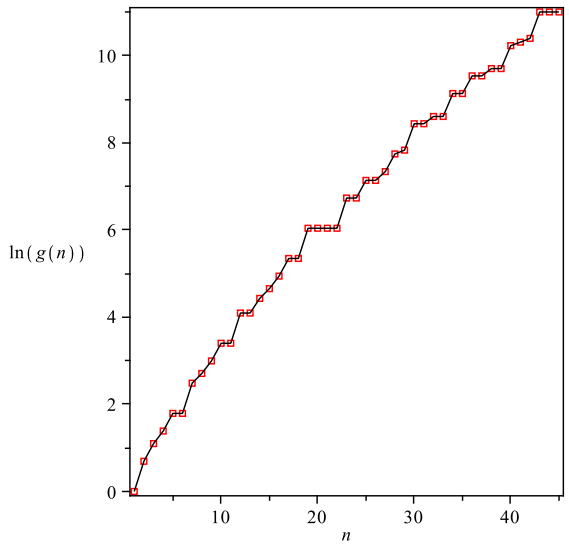
n	1	2	3	4	5	6	7	8	9	10
$g(n)$	1	2	3	4	6	6	12	15	20	30
n	11	12	13	14	15	16	17	18	19	20
$g(n)$	30	60	60	84	105	140	210	210	420	420
n	21	22	23	24	25	26	27	28	29	30
$g(n)$	420	420	840	840	1260	1260	1540	2310	2520	4620

This code is highly inefficient. For instance it takes 31.252 seconds to calculate $g(55)$, (which turned out to be the rather cool looking number $360,360$) and even when I made some simplifications it crashed my computer when I attempted to calculate $g(60)$. In fact the Landau function is quite a difficult function to deal with! For instance the most important and most famous unsolved conjecture in mathematics, the Reimann hypothesis, is equivalent to proving that it (the Landau function) is bounded in a certain way that it isn't too difficult to explain, but I won't go into explaining it!

The plots of the Landau function are below.



Once we get up to $n = 45$, these numbers really start jumping. For instance, $g(45) = 60060$, so the third plot is log plot.



Actually, Landau showed that

$$\lim_{n\to\infty} \frac{\ln g(n)}{\sqrt{n \ln n}} = 1.$$

EXERCISE 7.19. For $25 \le n \le 30$ verify that $\mathcal{S}_n$ contains an element of order $g(n)$, where $g(n)$ is as listed in the table.

SOLUTION

If we use the famous and fairly intuitive result which states that the sum of positive integers each greater than 1 is less than their product, we arrive at the following result.

THEOREM 7.50. *Suppose that m has a prime factorization*

$$m = p_1^{a_1} \cdots p_r^{a_r}$$

then the least integer n so that $\mathcal{S}_n$ contains an element of order m is

$$p_1^{a_1} + \cdots + p_r^{a_r}.$$

■

For instance, since $84 = 7 \cdot 3 \cdot 2^2$, $\mathcal{S}_{7+3+4} = \mathcal{S}_{14}$ contains an element of order 84, but $\mathcal{S}_{13}$ does not.

8 HOMOMORPHISMS

So far we have been very vague about what it means for two groups to have the same multiplication table. Let's talk about a few things that clearly should constitute "sameness" for two multiplication tables.

- If the table is written in a different order.
- If the names of the elements in the table are changed (in a consistent way).

To get a better handle on this let's consider three groups of order 4, namely the group $\mathbb{Z}_4$ (of course under addition) and $\{\{4\}\} = \{1, -1, i, -i\}$ (of course under multiplication), $\mathbb{Z}_2 \times \mathbb{Z}_2 = \mathbb{Z}_2^2$ under the operation $(+\ ,\ +)$, and $S(\blacksquare)$–the symmetries of a proper rectangle under composition.

$\mathbb{Z}_2^2$	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
$(0,0)$	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
$(1,0)$	$(1,0)$	$(0,0)$	$(1,1)$	$(0,1)$
$(0,1)$	$(0,1)$	$(1,1)$	$(0,0)$	$(1,0)$
$(1,1)$	$(1,1)$	$(0,1)$	$(1,0)$	$(0,0)$

$\{\{4\}\}$	1	-1	i	$-i$
1	1	-1	i	$-i$
-1	-1	1	$-i$	i
i	i	$-i$	-1	1
$-i$	$-i$	i	1	-1

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

$S(\blacksquare)$	$\mathbf{1}$	r_1	r_2	ϱ_π
$\mathbf{1}$	$\mathbf{1}$	r_1	r_2	ϱ_π
r_1	r_1	$\mathbf{1}$	ϱ_π	r_2
r_2	r_2	ϱ_π	$\mathbf{1}$	r_1
ϱ_π	ϱ_π	r_2	r_1	$\mathbf{1}$

In chapter four we argued informally that of these four groups can be split into two piles: the tables for $\mathbb{Z}_4$ and $\{\{4\}\}$ are "the same" and were put into one pile while the tables for the two groups $S(\blacksquare)$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$ are the same and were put in the opposite pile. We now introduce the formal definitions before moving on further.

8.1 Some Primitive Examples/Theory

Definition 8.1. Let $(M, \bullet)$ and $(N, \star)$ be algebraic systems. A function

$$\varphi\colon M \to N$$

is said to be *operation preserving* (o.p for short) if for all m_1 and m_2 in M we have that

$$\varphi(m_1 \bullet m_2) = \varphi(m_1) \star \varphi(m_2).$$

Remark 8.2. Note for an o.p. map φ we also have that

$$\varphi(m_1 \bullet m_2 \bullet \cdots \bullet m_{k-1} \bullet m_k) = \varphi(m_1) \star \varphi(m_2) \star \cdots \star \varphi(m_{k-1}) \star \varphi(m_k).^{\text{I}}$$

An operation preserving map between groups $(G, \bullet)$ and $(H, \star)$ gets a special name.

Definition 8.3.

- An operation preserving function between groups $\varphi\colon G \to H$ is called a *homomorphism (of groups)*.
- A bijective (i.e. invertible) homomorphism (of groups) is called an *isomorphism (of groups)*.
- If there exists an isomorphism $\varphi\colon G \to H$, then the groups G and H are said to be *isomorphic*.

Notation 8.4. If G and H are isomorphic, we use the notation $G \cong H$. If $\varphi\colon G \to H$ is an isomorphism, we write $\varphi\colon G \cong H$.

A function

$$\varphi\colon G \to H$$

is therefore an isomorphism if it satisfies the following three properties:

(1) φ is operation preserving.

(2) φ is surjective.

(3) φ is injective.

[I] We really should assume $\bullet$ is associative so these sorts of products without parenthesis make sense.

Of course, as for functions in general, we have two main ways of showing that a homomorphism is a bijection, and hence an isomorphism. The first is to show that it's both injective and surjective. The second is to find an explicit inverse.

REMARK 8.5. Therefore, at least in theory, if one needs to prove something about isomorphic groups the proof would look as follows:

> Suppose $G \cong H$ and G satisfies some property P (i.e. is Abelian, has an element of order 6, has order 52, etc).
> Then, there exists a function which preserves operations and is a bijection, $\varphi\colon G \to H$.
> ⋮
> (Some steps)
> ⋮
> Therefore H also satisfies property P.

For example, let's show this for P being the property "is abelian." That is suppose G is an abelian group and $G \cong H$, then we prove that H is abelian. I'm extra pedantic here and write every step one line at a time:

PROOF. Suppose G is abelian.

Suppose $G \cong H$.

Then there exists a function $\varphi\colon G \to H$ which is both operation preserving and a bijection.

Let h_1 and h_2 be two arbitrary elements of H. ⟨We want to show to show that $h_1h_2 = h_2h_1$⟩.

Since φ is a bijection it is in injective and surjective.

Therefore φ is surjective.

Since φ is surjective there exists elements g_1 and g_2 in G with $\varphi(g_1) = h_1$ and $\varphi(g_2) = h_2$.

Since G is abelian (by assumption) the elements g_1 and g_2 of G commute.

This means $g_1g_2 = g_2g_1$.

Since φ is single valued, $\varphi(g_1g_2) = \varphi(g_2g_1)$.

Since φ is operation preserving $\varphi(g_1)\varphi(g_2) = \varphi(g_2)\varphi(g_1)$

Since $\varphi(g_1) = h_1$ and $\varphi(g_2) = h_2$ this last step implies that $h_1h_2 = h_2h_1$.

Since h_1 and h_2 were arbitrary elements of H we have proven that H is (also) abelian. ■

Definition 8.6. These definitions shouldn't be your first priority, but we introduce them now because I don't know where else to stick them!

(1) A homomorphism from a group to itself is called an *endomorphism* of that group.

(2) A bijective endomorphism of a group is called an *automorphism* of that group.[I]

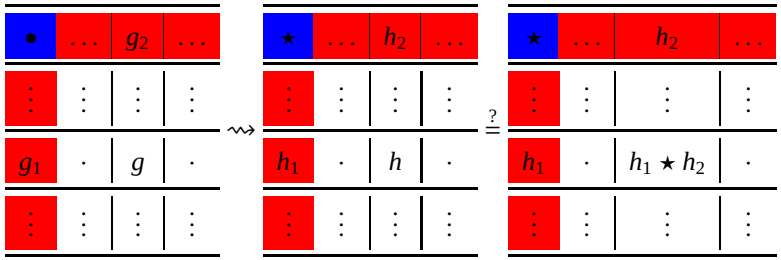
(3) An injective homomorphism of groups is called a *monomorphism*.[II]

(4) A surjective homomorphism of groups is called a *epimorphism*.[III]

For a function $\varphi\colon G \to H$ to be a homomorphism it really means that if we take a multiplication table for G, and everywhere we see the element g we replace it with the element $\varphi(g)$ we would have a table for H. Let g_1, g_2 be elements of G which has operation $\bullet$ and suppose that $g = g_1 \bullet g_2$ and denote

$$\varphi(g) = h, \quad \varphi(g_1) = h_1, \quad \varphi(g_2) = h_2$$

where h, h_1, and $h_2 \in H$. Then if we replace every entry of the table of G with the corresponding elements of H, then



Thus for the map φ to make the tables line up, we need that $h = h_1 \star h_2$. But $h = \varphi(g) = \varphi(g_1 \bullet g_2)$ and $h_1 = \varphi(g_1)$, $h_2 = \varphi(g_2)$, so we need

$$\varphi(g_1 \bullet g_2) = \varphi(g_1) \star \varphi(g_2)\,.$$

This equation is of course the requirement that the map φ is operation preserving, which we just defined. Therefore our definition of a homomorphism can be thought of as the requirement that φ makes the tables line up.

Example 8.7.

(1) Let's start with an example of a function which is not a homomorphism. Of course it is easier to show that a function is not a homomorphism than to show that it is, so this is why we start with this negative example. Let's take the function

$$\varphi\colon \mathbb{Z}_4 \to \mathbb{Z}_2 \times \mathbb{Z}_2$$

which we define explicitly as

$$\varphi([0]_4) = ([0]_2, [0]_2) \qquad \varphi([1]_4) = ([1]_2, [0]_2)$$
$$\varphi([2]_4) = ([0]_2, [1]_2) \qquad \varphi([3]_4) = ([1]_2, [1]_2).$$

So an automorphism is an isomorphism from a group to itself.

Mono, meaning 1, e.g. 1 to 1.

[I] Epi, as in epic, as in covering a lot of ground, as in onto ... ok, that was lame (epic failure ¨).

Then this particular function, φ, is *not* a homomorphism. This is illustrated as follows:

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

$\rightsquigarrow$

$\mathbb{Z}_2^2$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$
$([0]_2,[0]_2)$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$
$([1]_2,[0]_2)$	$([1]_2,[0]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$	$([0]_2,[0]_2)$
$([0]_2,[1]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$
$([1]_2,[1]_2)$	$([1]_2,[1]_2)$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$	$([1]_2,[1]_2)$

which is **not** a valid table. The below table corrects all of the "mistakes" in the above table.

$\mathbb{Z}_2^2$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$
$([0]_2,[0]_2)$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$
$([1]_2,[0]_2)$	$([1]_2,[0]_2)$	$([0]_2,[1]_2)$ (→ $([0]_2,[0]_2)$)	$([1]_2,[1]_2)$	$([0]_2,[0]_2)$ (→ $([0]_2,[1]_2)$)
$([0]_2,[1]_2)$	$([0]_2,[1]_2)$	$([1]_2,[1]_2)$	$([0]_2,[0]_2)$	$([1]_2,[0]_2)$
$([1]_2,[1]_2)$	$([1]_2,[1]_2)$	$([0]_2,[0]_2)$ (→ $([0]_2,[1]_2)$)	$([1]_2,[0]_2)$	$([1]_2,[1]_2)$ (→ $([0]_2,[0]_2)$)

Note that if the question simply asked whether or not φ were a homomorphism, we simply needed to find one pair that failed to test out. For instance,

$$\varphi([1]_4 + [1]_4) = \varphi([2]_4) = ([0]_2, [1]_2)$$

but

$$\varphi([1]_4) + \varphi([1]_4) = ([1]_2, [0]_2) + ([1]_2, [0]_2) = ([0]_2, [0]_2).$$

WARNING! 8.8. This only shows that this particular function φ is not a isomorphism. It does not show that there is no isomorphism between these two groups. Indeed we will see this to be the case, but at present just because we have shown that 1 bijection, φ, does not preserve operations does not prove that the other 23 bijections also do not preserve operations. To show this, without spending a lot of time, takes a bit of theory. We will come back to this later.

EXERCISE 8.1. Take a different bijection from $\mathbb{Z}_4 \to \mathbb{Z}_2 \times \mathbb{Z}_2$ than the one I just discussed and show that it is not an isomorphism.

SOLUTION

(2) For a positive instance, if we take the function

$$\varphi\colon \mathbb{Z}_4 \to \{\{4\}\}$$

given by

$$\varphi([0]_4) = 1 \qquad \varphi([1]_4) = i \qquad \varphi([2]_4) = -1 \qquad \varphi([3]_4) = -i$$

then we would take the table

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

$\rightsquigarrow$

$\{\{4\}\}$	1	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	1	$-i$	1	i
$-i$	$-i$	1	i	-1

which is a valid table for $\{\{4\}\}$. Thus φ is a homomorphism. Actually, since φ is a bijection φ is an isomorphism and we have proven that

$$\mathbb{Z}_4 \cong \{\{4\}\}.$$

Remark 8.9. Instead of verifying that the resulting table is valid, for me at least, I like to fill out the table myself and make sure it matches the resulting table. For instance, in our current situation I'd start with

$\{\{4\}\}$	1	i	-1	$-i$
1				
i				
-1				
$-i$				

and then fill out the table to get

$\{\{4\}\}$	1	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	1	$-i$	1	i
$-i$	$-i$	1	i	-1

.

Lastly, I compare this with what we got in the example and notice that the two are identical. We won't be doing many problems in this brute force way for long anyway, so whatever way you prefer to do these is fine!

Exercise 8.2. Find an explicit map

$$\mathbb{Z}_2^2 \to S(\blacksquare)$$

which is an isomorphism. (Please prove this map is an isomorphism but substituting in the table and verifying that the resulting table is valid.)

Solution

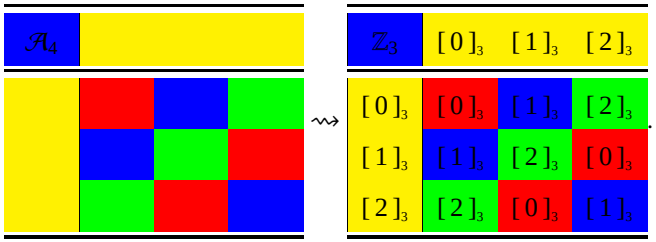
(3) Occasionally we can color a table in a certain way to see a homomorphism. For instance for group $\mathcal{A}_4$[II] with the dictionary

$$1 = (1) \qquad 2 = (1\,2)(3\,4) \qquad 3 = (1\,3)(2\,4)$$
$$4 = (1\,4)(2\,3) \qquad 5 = (1\,2\,3) \qquad 6 = (1\,3\,4)$$
$$7 = (2\,4\,3) \qquad 8 = (1\,4\,2) \qquad 9 = (1\,3\,2)$$
$$10 = (2\,3\,4) \qquad 11 = (1\,2\,4) \qquad 12 = (1\,4\,3)$$

we have the table

$\mathcal{A}_4$	1	2	3	4	5	6	7	8	9	10	11	12
1	1	2	3	4	5	6	7	8	9	10	11	12
2	2	1	4	3	7	8	5	6	12	11	10	9
3	3	4	1	2	8	7	6	5	10	9	12	11
4	4	3	2	1	6	5	8	7	11	12	9	10
5	5	6	7	8	9	10	11	12	1	2	3	4
6	6	5	8	7	11	12	9	10	4	3	2	1
7	7	8	5	6	12	11	10	9	2	1	4	3
8	8	7	6	5	10	9	12	11	3	4	1	2
9	9	10	11	12	1	2	3	4	5	6	7	8
10	10	9	12	11	3	4	1	2	8	7	6	5
11	11	12	9	10	4	3	2	1	6	5	8	7
12	12	11	10	9	2	1	4	3	7	8	5	6

Notice that the colors in this table can be thought of as a group which has the same multiplication table as $\mathbb{Z}_3$ after we replace red by $[0]_3$, blue by $[1]_3$, and green by $[2]_3$.



Thus we have a homomorphism (actually and epimorphism)

$$\varphi\colon \mathcal{A}_4 \to \mathbb{Z}_3$$

[II] Don't forget, as of now, the only reason that we that this set is a group at all is that we can eyeball it to see that it's closed. A boring reason indeed.

given by

$$\varphi((1)) = \varphi((1\,2)(3\,4)) = \varphi((1\,3)(2\,4)) = \varphi((1\,4)(2\,3)) = [0]_3$$
$$\varphi((1\,2\,3)) = \varphi((1\,3\,4)) = \varphi((2\,4\,3)) = \varphi((1\,4\,2)) = [1]_3$$
$$\varphi((1\,3\,2)) = \varphi((1\,4\,3)) = \varphi((2\,3\,4)) = \varphi((1\,2\,4)) = [2]_3 .$$

Exercise 8.3. Prove the map

$$\begin{pmatrix} \mathbf{1} & \rho & \rho^2 & \rho^3 & \mathbf{r} & \mathbf{r}\rho & \mathbf{r}\rho^2 & \mathbf{r}\rho^3 \\ \mathbf{1} & \rho & \rho^2 & \rho^3 & \mathbf{r}\rho & \mathbf{r}\rho^2 & \mathbf{r}\rho^3 & \mathbf{r} \end{pmatrix}$$

is an automorphism of $\mathcal{D}_4$.
Hint! Since you tell at a glance that this map is a bijection, $\mathcal{D}_4 \to \mathcal{D}_4$, you only need to prove that the map is operation preserving.
Hint! For now there is not quick way of doing this. Simply set up the tables and do this by brute force. We will come back to this later to present a much faster way of doing this!

Solution

8.2 Basic Theory

The methods of the previous examples are time consuming and not very theoretical. They give us a poor way of verifying something is a homomorphism without even giving us any insight of where these homomorphism come from. We start to explore examples which are a bit more theoretical in nature.

Theorem 8.10. *Suppose that*

$$\varphi\colon G \to H$$

is a homomorphism between groups G and H. Then for all $g \in G$

(1) $\varphi(e_G) = e_H$

(2) $\varphi(g^{-1}) = \varphi(g)^{-1}$

(3) $\varphi(g^n) = \varphi(g)^n$

Remark 8.11. In words, **(1)** says homomorphisms take the identity to the identity, part **(2)** says that homomorphisms preserve inverses, and **(3)** says that homomorphisms preserve powers. Of course part 2 is a special case of **(3)** (take $n = -1$), but we make it a separate part for emphasis.

Proof.

(1) Let $h = \varphi(e_G)$. Then

$$h = \varphi(e_G) = \varphi(e_G * e_G) = \varphi(e_G) \bullet \varphi(e_G) = h^2$$

but canceling h from both sides of $h = h^2$ shows us that $e_H = h = \varphi(e_G)$.

(2) We only need to show that $\varphi(g) \bullet \varphi(g^{-1}) = e_H$. But

$$\begin{aligned} \varphi(g) \bullet \varphi(g^{-1}) &= \varphi(g * g^{-1}) \\ &= \varphi(e_G) \\ &= e_H. \qquad \text{(By (1) just proven)} \end{aligned}$$

(3) In the case that $n > 0$ we can do this by induction on n. The case $n = 1$ holds by definition (since $g^1 = g$). If this holds for $n - 1$, then we have

$$\begin{aligned} \varphi(g^n) &= \varphi(g^{n-1}g) \\ &= \varphi(g^{n-1})\varphi(g) \\ &= \varphi(g)^{n-1}\varphi(g) \\ &= \varphi(g)^n. \end{aligned}$$

If $n < 0$, then we have $-n > 0$ so that

$$\begin{aligned} \varphi(g^n) &= \varphi((g^{-1})^{-n}) \\ &= \varphi(g^{-1})^{-n} \\ &= (\varphi(g)^{-1})^{-n} \qquad \text{(By (2) just proven)} \\ &= \varphi(g)^n. \end{aligned}$$

Thus the result holds in both cases. ∎

THEOREM 8.12. *Suppose that G and $\widehat{G}$ are groups and $H \leqslant G$ and $\widehat{H} \leqslant \widehat{G}$. Then for any group homomorphism $\varphi\colon G \to \widehat{G}$ we have*

(1) *The image of any subgroup is a subgroup, i.e. $\varphi(H) \leqslant \widehat{G}$*

(2) *The inverse image of any subgroup is a subgroup, i.e. $\varphi^{-1}(\widehat{H}) \leqslant G$.*

Two cases of this theorem are of particular importance. Namely

DEFINITION 8.13. Given a homomorphism $\varphi\colon G \to H$ we define

(1) The *image of* φ is the image of φ as a function, i.e. $\varphi(G)$. It is denoted as $\operatorname{Im}(\varphi)$.

(2) The *kernel of* φ is the inverse image of the set $\{e_H\}$. It is denoted as $\operatorname{Ker}(\varphi)$.

In symbols,

$$\begin{aligned} \operatorname{Im}(\varphi) &\stackrel{\text{def}}{=} \{\varphi(g) \in H : g \in G\} \\ \operatorname{Ker}(\varphi) &\stackrel{\text{def}}{=} \{g \in G : \varphi(g) = e_H\}. \end{aligned}$$

Example 8.14. The kernel of the homomorphism

$$\varphi\colon \mathcal{A}_4 \to \mathbb{Z}_3$$

given on page 375 has kernel which consists of the 3, 2 2–cycles and the identity:

$$\mathrm{Ker}\,(\varphi) = \{(1), (1\,2)(3\,4), (1\,3)(2\,4), (1\,4)(2\,3)\}.$$

Exercise 8.4. Suppose that $\varphi\colon G \to H$ is a homomorphism of groups. Prove that for any $g \in G$ we have

$$\varphi(\langle g \rangle) = \langle \varphi(g) \rangle .$$

Conclude that the image of any cyclic group under a homomorphism is a cyclic group.

Solution

Exercise 8.5. Consider the map defined by

$$\begin{aligned} \hat{}^3\colon \mathcal{U}(\mathbb{Z}_7) &\to \mathcal{U}(\mathbb{Z}_7) \\ : a &\mapsto a^3 \end{aligned}$$

(1) Prove that $\hat{}^3$ is a group homomorphism.

(2) Find $\mathrm{Ker}\left(\hat{}^3\right)$

(3) Find $\mathrm{Im}\left(\hat{}^3\right)$

(4) Verify that

$$|\mathcal{U}(\mathbb{Z}_7)| = \left|\mathrm{Ker}\left(\hat{}^3\right)\right| \cdot \left|\mathrm{Im}\left(\hat{}^3\right)\right|.$$

Solution

Exercise 8.6.

(1) Prove that

$$\begin{aligned} \varphi\colon \mathbb{Z}_{18} &\to \mathbb{Z}_{24} \\ : [a]_{18} &\mapsto 20 \cdot ([a]_{24}) . \end{aligned}$$

is well defined and is a homomorphism.

(2) Give the multiplication table (really the operation here is addition, but we still call it the multiplication table) for $\mathrm{Ker}\,(\varphi)$.

Solution

(3) Give the multiplication table for $\mathrm{Im}(\varphi)$.

(4) Verify that

$$|\mathbb{Z}_{18}| = |\mathrm{Ker}(\varphi)| \cdot |\mathrm{Im}(\varphi)|.$$

Proof. (**Proof of Theorem 8.12**)

(1) We use the one step subgroup test. This applies since $\varphi(H) \neq \emptyset$ since $\varphi(e_G)$ is in there.

OS1 Let $\widehat{h}, \widehat{h'} \in \varphi(H)$ be arbitrary. Then, by definition of $\varphi(H)$ we know there exists $h, h' \in H$ so that $\varphi(h) = \widehat{h}$, and $\varphi(h') = \widehat{h'}$. Since $H \leqslant G$ we know that H itself passes **OS1**, i.e. we have that $hh'^{-1} \in H$. Therefore, we have

$$h \cdot h'^{-1} \in H \Rightarrow \varphi(h \cdot h'^{-1}) \in \varphi(H)$$

by the general definition of $\varphi(H)$. But since

$$\varphi(h \cdot h'^{-1}) = \varphi(h) \cdot \varphi(h'^{-1}) = \varphi(h) \cdot \varphi(h')^{-1} \Rightarrow \varphi(h) \cdot \varphi(h')^{-1} = \widehat{h} \cdot \widehat{h'}^{-1} \in \varphi(H)$$

as well. As $\widehat{h}$ and $\widehat{h'}$ were arbitrary elements of $\varphi(H)$, we have proved that $\varphi(H)$ passes **OS1**.

(2) For starters, we know that $\varphi^{-1}(\widehat{H}) \neq \emptyset$ since

$$\varphi(e_G) = e_{\widehat{G}} \in \widehat{H} \Rightarrow e_G \in \varphi^{-1}(\widehat{H}).$$

Therefore, we can apply one of the subgroup tests, which, for variety, we use the two step subgroup test. Let $h, \dot{h} \in \varphi^{-1}(\widehat{H})$ be arbitrary.

TS1 We show that $\varphi^{-1}(\widehat{H})$ is closed under inversion. Since $\widehat{H}$ is a subgroup and $\varphi(h) \in \widehat{H}$ by definition of an inverse image, we know that

$$\varphi(h^{-1}) = \varphi(h)^{-1} \in \widehat{H} \Rightarrow h^{-1} \in \varphi^{-1}(\widehat{H})$$

since we are assuming that $\widehat{H} \leqslant \widehat{G}$. Therefore

$$h^{-1} \in \varphi^{-1}(\widehat{H})$$

which completes the proof that $\varphi^{-1}(\widehat{H})$ passes **TS1**.

TS2 We now prove that $\varphi^{-1}(\widehat{H})$ is closed under multiplication. Since $\varphi(h)$ and $\varphi(\dot{h})$ are both elements of $\widehat{H}$ and since $\widehat{H} \leqslant \widehat{G}$, we must have that the product, $\varphi(h)\varphi(\dot{h})$ is also an element of the subgroup (of $\widehat{G}$) $\widehat{H}$ by the closure. But by definition of a homomorphism, we know that this product equals $\varphi(h\dot{h})$. Therefore

$$\varphi(h\dot{h}) \in \widehat{H} \Rightarrow h\dot{h} \in \varphi^{-1}(\widehat{H}).$$

Since h and $\dot{h}$ were arbitrary, we have proven that $\varphi^{-1}(\widehat{H})$ passes **TS2**. ■

Exercise 8.7. For p a prime, prove that every homomorphism $G \to \mathbb{Z}_p$ is either surjective or sends every element of G to $[0]_p$.
Hint! Exercise/Theorem 6.26 and the above.

SOLUTION

Now the image of a homomorphism, like the image of any function, detects how far a homomorphism is from being surjective in an obvious way; namely that

$$\varphi\colon G \to H$$

is surjective, iff $\mathrm{Im}(\varphi) = H$. This is a fact that is true, just about by definition, for all functions, even between sets which are not groups. On the other hand, the kernel of a homomorphism is something special since it involves the identity, which an arbitrary set has no notion of. But we are very lucky to have the kernel in our group theory quest, however, because it has the property that it can detect how far away a homomorphism is from being **injective.** Indeed, we will see that a function is injective iff the kernel contains a single element. Thus, in a precise sense, the smaller the kernel is the closer the function is to being injective, while the larger the image is the closer the function is to being surjective.

Obviously, if a homomorphism is injective, it has trivial kernel, meaning the kernel has only 1 element, the identity. This is because the kernel is the collection of guys in the first group, the domain of the homomorphism, which get mapped to the identity of the second group. Therefore, if the kernel has more than 1 element in it (i.e. is non-trivial), then more than one element of the domain gets mapped to the identity of the second group, and so the homomorphism could not possibly be injective. The coolest thing about kernels is that the opposite is also true. This is a powerful statement; it says that two elements of the first group go to one element of the second iff two elements of the first group go to the identity of the second group. It is a very important theorem that should be internalized in the sense that whenever we want to prove that a certain homomorphism is injective, we immediately try to prove that the kernel is trivial.

THEOREM 8.15. *A homomorphism is injective iff its kernel is trivial.*

PROOF. We have already proven that monomorphisms have trivial kernel. Suppose that $\varphi\colon G \to H$ is a homomorphism with trivial kernel, i.e. a homomorphism which satisfies

$$\varphi(g) = e_H \Rightarrow g = e_G.$$

Now suppose that $\varphi(g) = \varphi(\dot{g})$ for two elements $g, \dot{g} \in G$. We need to show that $g = \dot{g}$. We have

$$\begin{aligned}
\varphi(g) = \varphi(\dot{g}) &\Rightarrow \varphi(g)\varphi(\dot{g})^{-1} = \varphi(\dot{g})\varphi(\dot{g})^{-1} && \text{(Mult on rt by } \varphi(\dot{g})^{-1}\text{)}\\
&\Rightarrow \varphi(g)\varphi(\dot{g})^{-1} = e_H && \text{(Cancel element w/inv)}\\
&\Rightarrow \varphi(g)\varphi(\dot{g}^{-1}) = e_H && \text{(By 8.10(2))}\\
&\Rightarrow \varphi(g\dot{g}^{-1}) = e_H && \text{(By def'n of a hom)}\\
&\Rightarrow g\dot{g}^{-1} \in \mathrm{Ker}(\varphi) && \text{(By def'n of the kernel)}\\
&\Rightarrow g\dot{g}^{-1} = e_G && \text{(Since } e_G \text{ is the only ele of ker)}\\
&\Rightarrow g = \dot{g}
\end{aligned}$$

as desired. ∎

In symbols, this theorem says

$$\varphi\colon G \to H \text{ is injective } \Leftrightarrow \mathrm{Ker}(\varphi) = \{e_G\}.$$

REMARK 8.16. We will see that in general a homomorphism is m to 1 (i.e. the set of all elements of the first set which map to any given element of the image of the map has m elements in it.) iff the kernel of the homomorphism has m elements.

Recall for subsets K and S of a group G,

$$KS \overset{\text{def}}{=} \{ks : k \in K, s \in S\}.$$

Example /Theorem 8.17. For any $S \subseteq G$ and homomorphism φ with domain G and kernel K, we have

$$\varphi^{-1}(\varphi(S)) = KS.$$

Proof. We are attempting to prove the equality of sets, so we show mutual containment.

$\subseteq$ Let $g \in \varphi^{-1}(\varphi(S))$ be arbitrary. Then $\varphi(g) \in \varphi(S)$. This means there exists an $s \in S$ so that

$$\begin{aligned}
\varphi(g) = \varphi(s) &\Rightarrow \varphi(g)\varphi(s)^{-1} = \varphi(s)\varphi(s)^{-1} = e \\
&\Rightarrow \varphi(g)\varphi(s^{-1}) = e \\
&\Rightarrow \varphi(gs^{-1}) = e \\
&\Rightarrow gs^{-1} \in \operatorname{Ker}(\varphi) \\
&\Rightarrow gs^{-1} = k \text{ for some } k \in \operatorname{Ker}(\varphi) \\
&\Rightarrow g = ks \\
&\Rightarrow g \in KS.
\end{aligned}$$

$\supseteq$ All arrows can be reversed to give the opposite inclusion. ∎

Exercise 8.8. Suppose that $K \leqslant G$ and that $\varphi\colon G \to H$ is a group homomorphism.

(1) Show that

$$\operatorname{Ker}(\varphi|_K) = \operatorname{Ker}(\varphi) \cap K.$$

(2) Suppose further that $K \cap \operatorname{Ker}(\varphi)$ is trivial. Then show that $K \cong \varphi(K)$.

SOLUTION

Exercise 8.9. Consider the function f given by $f(x) = x^2$. Show that $f(a \cdot b) = f(a) \cdot f(b)$ and maps only 0 to 0, but it is not injective. Why doesn't this contradict Theorem 8.15?

SOLUTION

Exercise 8.10.

(1) Prove that if φ is an isomorphism, then φ^{-1} is as well. (So you must prove that φ^{-1} is operation preserving and bijective.)

(2) Prove that the composition of two homomorphisms (when defined) is a homomorphism.

(3) Prove that the set of all isomorphisms from a group G to itself forms a group (under composition) called the *automorphism group* of G. This group is denoted $\operatorname{Aut}(G)$.

Hint! These dudes are functions. As such, $\operatorname{Aut}(G) \subseteq \mathcal{S}_G$. Since $\mathcal{S}_G$ is a group (the bijections of any set form a group) you can apply a subgroup test to $\operatorname{Aut}(G)$: i.e. you simply need to show that if φ_1 and φ_2 are isomorphisms $G \to G$, then $\varphi_1 \circ \varphi_2$ is and φ_1^{-1} is ... wait, didn't you just do that?

SOLUTION

(4) Prove that the relation $G \cong H$ defines an equivalence relation on the class of all groups.
Hint! You're basically already done. For a template see the proof of Theorem 2.102 back on page 69.

Hint! We have proven these a few times before in specific cases. However, the proofs done were always done in a way that didn't use anything specific about the situation and as such carry over to the present situation.

Here are some more basic points and observations:

(1) Obviously if we have a homomorphism $\varphi\colon G \to \widehat{G}$ and subgroups $H \leqslant G$ and $\widehat{H} \leqslant \widehat{G}$ with $\varphi(H) \subseteq \widehat{H}$, then the restricted function

$$\restriction{\varphi}{H}^{\widehat{H}}\colon H \to \widehat{H}$$

is still a group homomorphism since the only quantifiers involved in the definition of a group homomorphism are "for alls." In particular, since we just showed that $\varphi(H) \leqslant \widehat{H}$ we have a epimorphisms

$$H \to \varphi(H) \quad \text{and} \quad G \to \operatorname{Im}(\varphi).$$

(2) If $\varphi\colon G \to \widehat{G}$ is a monomorphism (i.e. an injective homomorphism) of groups, then $G \cong \operatorname{Im}(\varphi)$. That is a group is isomorphic to its image under a monomorphism. This is because the corestriction

$$\varphi\colon G \to \operatorname{Im}(\varphi)$$

is still a homomorphism which is injective. Moreover, every function surjects to its image. Therefore, the corestriction is a bijection and a homomorphism, i.e. an isomorphism. In this way we can think of G as being a subgroup of $\widehat{G}$ whenever there exists a monomorphism $G \to \widehat{G}$.

Notation 8.18. Let G and A be groups. Define

$$\hom(G, A) \stackrel{\text{def}}{=} \{\varphi\colon G \to A \mid \varphi \text{ is a homomorphism}\},$$

the set of all homomorphisms $G \to A$.

8.3 Millions and Millions of Examples

In this section we give a ton of examples. We won't however show you how to "generate" examples until Chapter 10. Therefore most of the examples in this section are of the form: the following function is a homomorphism since blah blah blah.

Example 8.19.

(1) Since

$$\mathbf{1}_G(g_1 \star g_2) = g_1 \star g_2 = \mathbf{1}_G(g_1) \star \mathbf{1}_G(g_2)$$

we see that the identity map is a homomorphism from any group to G to itself. Moreover, since the identity map is invertible, the identity map is an isomorphism. In particular any group is isomorphic to itself.

(2) Similar to the previous example, for any subgroup $H \leqslant G$, the inclusion

$$\iota : H \to G$$

is a homomorphism ($\iota(h) = h$ for all $h \in H$) which is injective (hence a monomorphism).

(3) The trivial homomorphism, denoted $\mathbf{0} : G \to H$, is defined to be the function which sends every element of G to the identity of H (so $\mathbf{0}(g) = e_H$ for every $g \in G$). Indeed it really is a homomorphism since

$$\mathbf{0}(g_1 \star g_2) = e_H = e_H \bullet e_H = \mathbf{0}(g_1) \bullet \mathbf{0}(g_2).$$

EXERCISE 8.11. For a group homomorphism, $\varphi : G \to H$, prove that TFAE

(1) $\varphi = \mathbf{0}$;

(2) $\operatorname{Ker}(\varphi) = G$;

(3) $\operatorname{Im}(\varphi) = \{e_H\}$.

SOLUTION

(4) The function $\exp : \mathbb{R} \to \mathbb{R}^+$ (the operation in the first set is $+$ and in the second it is $\cdot$) given by $\exp(x) = e^x$ from calculus is a homomorphism of groups (the first has addition as its operation, the second has multiplication) since

$$\exp(x + y) = e^{x+y} = e^x e^y = \exp(x) \cdot \exp(y).$$

Moreover it is an isomorphism since it's a bijection (with inverse $\ln$, the natural log).

(5) Similar to the last example, the function

$$\begin{aligned} \operatorname{Exp} &: \mathbb{R} \to \mathbb{S}^2 \\ &: r \mapsto \exp(2\pi i r) \end{aligned}$$

is a homomorphism. But unlike the previous problem, it is not injective. In fact, the kernel of Exp, which by definition is the set of real numbers r so that $\exp(2\pi i r) = 1$, is the set $\mathbb{Z}$. This is due to the fact that $\exp(2\pi i r) = \cos 2\pi r + i \sin 2\pi r$ which equals 1 iff r is an integer. In symbols

$$\operatorname{Ker}(\operatorname{Exp}) = \mathbb{Z}.$$

(6) In basic math courses at the beginning of ones undergraduate career, poor, hapless students think that everything is a homomorphism. Common *very* false things students believe include:

$$\cancel{(x+y)^2 = x^2 + y^2}$$
$$\cancel{\sqrt{x+y} = \sqrt{x} + \sqrt{y}}$$
$$\cancel{\sin(x+y) = \sin x + \sin y.}$$

Of course none of these statements are true. In our language, none of the functions mentioned (the square function, the square root function, the sin function) are operation preserving.

(7) Whenever A is an abelian group (written in multiplicative notation) and n is an integer, the function

$$\hat{}^n : a \mapsto a^n$$

is a group homomorphism. This is because for abelian groups, we have $(ab)^n = a^n b^n$ for every pair of elements, c.f. Theorem 5.32 **(5)**. In particular, for $n = -1$ we have a homomorphism $a \mapsto a^{-1}$. If we write A additively, we are saying that

$$n \cdot (a + b) = na + nb$$

which of course, resembles the distributive property. Note that the kernel of the map $\hat{}^n$ is by definition

$$\operatorname{Ker}(\hat{}^n) \stackrel{\text{def}}{=} \{a \in A : a^n = e_A\}$$

which is the very definition of the set $A[\,n\,]$ given on page 264.

(8) The square root function

$$\begin{aligned} \sqrt{} &: \mathbb{R}^+ \to \mathbb{R}^+ \\ &: x \mapsto \sqrt{x} \end{aligned}$$

is an isomorphism of groups. (Here $\mathbb{R}^+$ is the multiplicative group of positive real numbers) It is the inverse to

$$\begin{aligned} \hat{}^2 &: \mathbb{R}^+ \to \mathbb{R}^+ \\ &: x \mapsto x^2. \end{aligned}$$

Surprisingly, there is no reasonable way to define $\sqrt{}$ for all of $\mathbb{C}^\times$ so this still holds. For instance, it seems reasonable to define $\sqrt{-1}$ to be i (as opposed to $-i$) but then

$$-1 = (i) \cdot (i) = \sqrt{-1} \cdot \sqrt{-1} \neq \sqrt{(-1) \cdot (-1)} = \sqrt{1} = 1.$$

The same problems occurs if we define $\sqrt{-1}$ to be $-i$.

Exercise 8.12. Prove that TFAE

Solution

(1) The map

$$\begin{aligned} \hat{}^{-1} &: G \to G \\ &: g \mapsto g^{-1} \end{aligned}$$

is a group homomorphism.

(2) The map

$$\begin{aligned} \hat{}^2 &: G \to G \\ &: g \mapsto g^2 \end{aligned}$$

is a group homomorphism.

(3) For any integer n, the map

$$\begin{aligned} \hat{}^n &: G \to G \\ &: g \mapsto g^n \end{aligned}$$

is a group homomorphism.

(4) G is abelian.

EXERCISE 8.13. Show that for any non-zero integers n and m, that the groups $n\mathbb{Z} \cong m\mathbb{Z}$.

SOLUTION

(9) A very important example of an isomorphism for physicists says that $\mathbb{S}^3 \cong \mathbf{SU}_2(\mathbb{C})$ where $\mathbf{SU}_2(\mathbb{C})$ is the special unitary group which was defined back in Exercise 6.5 on page 323. Recall that the $\mathbb{C} \times \mathbb{C}$ notation for a quaternion wrote a quaternion $\mathsf{q} = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ as the ordered pair of complex numbers $(a + bi, c + di) \in \mathbb{C}^2$. Recall in that same exercise, you prove that

$$\mathbf{SU}_2(\mathbb{C}) = \left\{ \begin{pmatrix} \mathsf{z} & \mathsf{w} \\ -\overline{\mathsf{w}} & \overline{\mathsf{z}} \end{pmatrix} : \mathsf{z}, \mathsf{w} \in \mathbb{C}, \|\mathsf{z}\|^2 + \|\mathsf{w}\|^2 = 1 \right\}.$$

First we define φ as the function which takes the matrix

$$\begin{pmatrix} \mathsf{z} & \mathsf{w} \\ -\overline{\mathsf{w}} & \overline{\mathsf{z}} \end{pmatrix}$$

in $\mathbf{SU}_2(\mathbb{C})$ to the $\mathbb{C} \times \mathbb{C}$ quaternion (z, w) (which in $\mathbb{R}^4$ notation is the same as the quaternion $\mathsf{z} + \mathsf{w}\mathbf{j}$).

- (φ is a bijection.) Every quaternion can be uniquely written in $\mathbb{C} \times \mathbb{C}$ notation. For a quaternion (z, w) written in $\mathbb{C} \times \mathbb{C}$ notation,

  $$\|(\mathsf{z}, \mathsf{w})\|^2 = \|\mathsf{z}\|^2 + \|\mathsf{w}\|^2.$$

  Thus, in $\mathbb{C} \times \mathbb{C}$ notation, $\mathbb{S}^3$ is the set

  $$\{(\mathsf{z}, \mathsf{w}) : \|\mathsf{z}\|^2 + \|\mathsf{w}\|^2 = 1\}.$$

  This means that for any $(\mathsf{z}, \mathsf{w}) \in \mathbb{S}^3$,

  $$\begin{pmatrix} \mathsf{z} & \mathsf{w} \\ -\overline{\mathsf{w}} & \overline{\mathsf{z}} \end{pmatrix}$$

  is the unique matrix which maps to it. Hence φ is a bijection.
- (φ is a operation preserving.) Let

  $$\mathbf{A}_1 = \begin{pmatrix} \mathsf{z}_1 & \mathsf{w}_1 \\ -\overline{\mathsf{w}}_1 & \overline{\mathsf{z}}_1 \end{pmatrix} \quad \text{and} \quad \mathbf{A}_2 = \begin{pmatrix} \mathsf{z}_2 & \mathsf{w}_2 \\ -\overline{\mathsf{w}}_2 & \overline{\mathsf{z}}_2 \end{pmatrix}$$

  be two arbitrary matrices in $\mathbf{SU}_2(\mathbb{C})$. Then

  $$\begin{aligned} \mathbf{A}_1 \cdot \mathbf{A}_2 &= \begin{pmatrix} \mathsf{z}_1 & \mathsf{w}_1 \\ -\overline{\mathsf{w}}_1 & \overline{\mathsf{z}}_1 \end{pmatrix} \cdot \begin{pmatrix} \mathsf{z}_2 & \mathsf{w}_2 \\ -\overline{\mathsf{w}}_2 & \overline{\mathsf{z}}_2 \end{pmatrix} \\ &= \begin{pmatrix} \mathsf{z}_1\mathsf{z}_2 - \mathsf{w}_1\overline{\mathsf{w}}_2 & \mathsf{z}_1\mathsf{w}_2 + \mathsf{w}_1\overline{\mathsf{z}}_2 \\ \times\times\times\times & \times\times\times\times \end{pmatrix} \end{aligned}$$

(we have not bothered to calclulate the two $\times\times\times\times$ terms in the above matrix since they are not relevant in order to calculate $\varphi(\mathbf{A}_1 \cdot \mathbf{A}_2)$.) Therefore, we have that

$$\varphi(\mathbf{A}_1 \cdot \mathbf{A}_2) = (z_1 z_2 - w_1 \overline{w}_2, z_1 w_2 + w_1 \overline{z}_2).$$

Meanwhile, $\varphi(\mathbf{A}_1) = (z_1, w_1)$ and $\varphi(\mathbf{A}_2) = (z_2, w_2)$ and by the formula for $\mathbb{C} \times \mathbb{C}$ multiplication, proved in Theorem 3.95 on page 120, we know that

$$(z_1, w_1) * (z_2, w_2) = (z_1 z_2 - w_1 \overline{w}_2, z_1 w_2 + w_1 \overline{z}_2)$$

as well. In summary, we have proved that

$$\varphi(\mathbf{A}_1 \cdot \mathbf{A}_2) = \varphi(\mathbf{A}_1) * \varphi(\mathbf{A}_2)$$

(the operation on the LHS is matrix multiplication, while the operation on the RHS is quaternion multiplication) which means that φ is operation preserving.

In conclusion, φ is an isomorphism of groups, $\mathbf{SU}_2(\mathbb{C}) \cong \mathbb{S}^3$.

(10) The determinant function

$$\det\colon \mathbf{M}_n(R) \to R$$

is operation preserving whenever R is a commutative ring (e.g. $R = \mathbb{Z}_n$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$) since we have for any two matrices $\mathbf{A}$ and $\mathbf{B}$ that

$$\det(\mathbf{AB}) = \det\mathbf{A} \cdot \det\mathbf{B}$$

by Theorem 3.205 back on page 189. (So det takes matrix multiplication to the multiplication in R.) Of course the determinant of an invertible matrix is an invertible element of R, so the image of this function is indeed contained inside of $\mathcal{U}(R)$ (The set of invertible elements of R). Therefore

$$\det\colon \mathbf{GL}_n(R) \to \mathcal{U}(R)$$

is a homomorphism of groups. The kernel of this homomorphism, since the identity of $\mathcal{U}(R)$ is 1, is the set of matrices whose determinant is 1. But, by its very definition, this is the set $\mathbf{SL}_n(R)$.

(11) Similar to the last example, for all positive integers n, the map

$$\det\colon \mathcal{D}_n \to \mathbb{Z}_2$$

is a homomorphism. It takes all rotations to 1 and all reflections to -1. Hence the kernel of this homomorphism consists of the set of rotations, $\langle \rho \rangle$, inside of $\mathcal{D}_n$.

(12) In Appendix C, actually in Definition C.33 on page 779, we define the map

$$\begin{aligned}\mathbf{R}\colon \mathbb{S}^3 &\to \mathbf{SO}_3(\mathbb{R})\\ : \cos\frac{\theta}{2} + \sin\frac{\theta}{2}\widehat{\mathbf{u}} &\mapsto \varrho_{\theta,\widehat{\mathbf{u}}}.\end{aligned}$$

Moreover in that appendix, (Theorem C.41 coming up on page 782) we end proving that this map is an epimorphism with kernel $\{\pm 1\}$.

For a set $X \subseteq \mathbb{R}^3$, $S(X)$, we define $B(X)$ as the set of quaternions in $\mathbb{S}^3$ which correspond to an (orientation preserving) symmetry of X. In more detail

$$B(X) = \{\cos\theta + \sin\theta\,\widehat{\mathbf{u}} : \text{Rotation by angle } 2\theta \text{ about the axis } \widehat{\mathbf{u}} \text{ is a symmetry of } X\}.$$

Notice that since

$$\varrho_{2\theta,\widehat{\mathbf{u}}} = \varrho_{-2\theta,-\widehat{\mathbf{u}}} = \varrho_{2\pi-2\theta,-\widehat{\mathbf{u}}}$$

we know that

$$\cos\theta + \sin\theta\widehat{\mathbf{u}} \in B(X) \Leftrightarrow \cos(\pi-\theta) - \sin(\pi-\theta)\widehat{\mathbf{u}} = -\cos\theta - \sin\theta\widehat{\mathbf{u}} \in B(X)$$

i.e.

$$\mathsf{q} \in B(X) \Leftrightarrow -\mathsf{q} \in B(X).$$

DEFINITION 8.20. $B(X)$ is called the *binary X group* or the *double* of the symmetry group of X, $S(X)$.

Note that more formally, $B(X)$ can is the set $\mathbf{R}^{-1}(S(X))$, i.e. the set

$$\{\mathsf{q} \in \mathbb{S}^3 : \mathbf{R}_{\mathsf{q}} \in S(X)\}.$$

This last way of thinking of $B(X)$ immediately implies that $B(X)$ is a group (via Theorem 8.12**(2)**, which says the inverse image under an homomorphism, which $\mathbf{R}$ is a perfectly good example of, of any subgroup of the codomain, which $S(X)$ is a perfectly good example of, is a subgroup of the domain). Moreover the (restricted map) $\mathbf{R}\colon B(X) \to S(X)$ is surjective and is still 2 to 1 (these last two facts have nothing to do with group theory). In particular we have

$$|B(X)| = 2 \cdot |S(X)|.$$

We've already seen one instance of this: the binary polygon groups defined in Chapter 5 on page 304, $B\mathcal{D}_n$. In our new notation, these groups would be denoted $B(\{\{n\}\})$, the double of the dihedral groups.

In addition to the groups $B\mathcal{D}_n$, there are 3 more special cases of particular importance:

DEFINITION 8.21. The

(1) Binary Tetrahedral group, $B\left(\blacktriangleright^3\right) \stackrel{\text{def}}{=} \mathbf{R}^{-1}\left(S\left(\blacktriangleright^3\right)\right)$;

(2) Binary Octahedral group, $B\left(\blacklozenge^3\right) \stackrel{\text{def}}{=} \mathbf{R}^{-1}\left(S\left(\blacklozenge^3\right)\right)$;

(3) Binary Icosahedral group, $B\left(\bigstar^3\right) \stackrel{\text{def}}{=} \mathbf{R}^{-1}\left(S\left(\bigstar^3\right)\right)$.

For instance, for the standard cube, rotation about the top face by 90° is a symmetry. The axis of this rotation is $\mathbf{k}$ and so

$$\cos\frac{\pi}{4} + \sin\frac{\pi}{4}\mathbf{k} = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{k} \in B\left(\blacklozenge^3\right).$$

{Insert Image}

$$\text{Symmetry groups of order } n \rightsquigarrow \text{Subgroups of } \mathbb{S}^3 \text{ of order } 2n.$$

(13) (Calculus example) For a function from $F\colon \mathbb{R}^n \to \mathbb{R}^m$ the *Jacobian* matrix of F (also known as the matrix of partials), denoted dF is the m by n matrix whose i, j entry is given by

$$\frac{\partial F_i}{\partial x_j}.$$

For instance the function

$$F(x, y) = (x^2y^2, xy, x^2 + y^2)$$

is

$$\begin{pmatrix} 2xy^2 & 2x^2y \\ y & x \\ 2x & 2y \end{pmatrix}.$$

If we were to plug in a point in $\mathbf{p} \in \mathbb{R}^n$ into this matrix, we denote the resulting real matrix by $dF_{\mathbf{p}}$. Now the set of functions from $\mathbb{R}^n \to \mathbb{R}^n$, like the set of functions from any set to itself, forms a monoid (hence an algebraic system). The chain rule from a calculus 3 course states that for two such functions F and G we have

$$d(F \circ G)_{\mathbf{p}} = d(F)_{G(\mathbf{p})} \cdot d(G)_{\mathbf{p}}.$$

In the special case that $n = 1$, this of course is just the usual chain rule one sees in calculus 1. This is very close to the statement that for a fixed such $\mathbf{p}$ the function $F \mapsto d_{\mathbf{p}}F$ is operation preserving (the first operation is composition of functions and the second operation is multiplication of matrices). Indeed it is an operation preserving map when one restricts to the monoid of functions which take the point $\mathbf{p}$ to itself. For instance the functions

$$F(x, y) = (x^2 + y^2 - y, xy + x - y) \quad \text{and} \quad G(x, y) = (x^2y + x - y, xy^2)$$

each take the point $(1, 1)$ to itself. Meanwhile

$$dF_{(1,1)} = \begin{pmatrix} 2 & 1 \\ 2 & 0 \end{pmatrix} \quad \text{and} \quad dG_{(1,1)} = \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}$$

and their product is the matrix

$$dF_{(1,1)} \cdot dG_{(1,1)} = \begin{pmatrix} 7 & 2 \\ 6 & 0 \end{pmatrix}$$

This checks since the composition $F \circ G$ takes a point (x, y) to

$$((x^2y + x - y)^2 + (xy^2)^2 - (xy), (x^2y + x - y)(xy^2) + (x^2y + x - y) - (xy^2))$$

and this can be checked to have the same Jacobian at the point $(1, 1)$. In many ways, this operation preserving map is the basis of many aspects of calculus. It says after shifting a function around functions "look" linear in such a way that the composition of the functions corresponds to (the much much easier!) composition of linear functions (i.e. matrix multiplication).

(14) We define a homomorphism $\mathbf{F} \colon \mathcal{S}_n \to \mathbf{GL}_n(\mathbb{Z})$[I] by sending a permutation σ to the linear transformation which permutes the standard basis vectors according to σ, that is

$$\mathbf{F}_\sigma(\mathbf{e}_i) = \mathbf{e}_{\sigma(i)}.$$

In other words, this is the matrix that is a permutation of the identity matrix obtained by permuting the rows according to σ.

To show that $\mathbf{F}$ is a homomorphism we must show that $\mathbf{F}_{\sigma\tau} = \mathbf{F}_\sigma \circ \mathbf{F}_\tau$ for all permutations σ and τ. Since the $\mathbf{F}$'s are linear maps, we simply need to show that both sides do the same thing to the standard basis vectors $\mathbf{e}_i$. Indeed

$$\begin{aligned} \mathbf{F}_{\sigma\tau}(\mathbf{e}_i) &= \mathbf{e}_{\sigma(\tau(i))} \\ &= \mathbf{F}_\sigma(\mathbf{e}_{\tau(i)}) \\ &= \mathbf{F}_\sigma(\mathbf{F}_\tau(\mathbf{e}_i)) \\ &= \big(\mathbf{F}_\sigma \circ \mathbf{F}_\tau\big)(\mathbf{e}_i). \end{aligned}$$

[I] As usual in functions like this, we write the image of an element σ as $\mathbf{F}_\sigma$ instead of $\mathbf{F}(\sigma)$ so it's a bit easier to see.

For instance, let's write the image of $\mathbf{F}$ when applied to $\mathcal{S}_4$.

$$\mathbf{F}_{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \mathbf{F}_{(12)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \mathbf{F}_{(23)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{F}_{(34)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \mathbf{F}_{(123)} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \mathbf{F}_{(124)} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\mathbf{F}_{(234)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \mathbf{F}_{(12)(34)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \mathbf{F}_{(132)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{F}_{(243)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \mathbf{F}_{(1234)} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \mathbf{F}_{(1342)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\mathbf{F}_{(1243)} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \mathbf{F}_{(1432)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \mathbf{F}_{(1342)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\mathbf{F}_{(13)} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \mathbf{F}_{(24)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \mathbf{F}_{(13)(24)} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\mathbf{F}_{(14)} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \mathbf{F}_{(1423)} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \mathbf{F}_{(14)(23)} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Matrices in the image of $\mathbf{F}$ are known as *permutation matrices.* Notice that $\mathbf{F}$ is injective since if $\sigma \in \mathrm{Ker}\,(\mathbf{F})$, then $\mathbf{F}_\sigma = \mathbf{1}$, the identity of $\mathbf{GL}_n(\mathbb{Z})$ (here we are attempting to use Theorem 8.15. Therefore we are trying to show that $\sigma = (1)$). In particular,

$$\mathbf{F}_\sigma(\mathbf{e}_i) = \mathbf{e}_{\sigma(i)} = \mathbf{e}_i$$

for all standard basis vectors $\mathbf{e}_i$. But this means that $\sigma(i) = i$ for all i and therefore $\sigma = (1)$, the identity of $\mathcal{S}_n$.

EXERCISE 8.14. Prove for permutation σ, $\mathbf{F}_\sigma^{-1} = \mathbf{F}_\sigma^\top$, i.e that $\mathbf{F}_\sigma$ is an orthogonal matrix.

SOLUTION

Exercise 8.15. Let $H \subseteq \mathbf{GL}_n(\mathbb{Z})$ be the set of matrices with exactly one 1 in every row and column (and nothing else).

Solution

(1) Prove that H is the set of permutation matrices.

(2) Prove that $H \cong \mathcal{S}_n$.

Exercise 8.16. For a permutation σ, prove that $\operatorname{tr}(\mathbf{F}_\sigma)$ is equal to the number of fixed points of σ.

Solution

Exercise 8.17. Let $\mathbf{A}$ be a square matrix and σ be a permutation. Prove that (assuming that these products make sense)

Solution

(1) $(\mathbf{A} \cdot \mathbf{F}_\sigma)(i, j) = \mathbf{A}(i, \sigma(j))$

(2) $(\mathbf{F}_\sigma \cdot \mathbf{A})(i, j) = \mathbf{A}(\sigma^{-1}(i), j)$

(3) $\left(\mathbf{F}_\sigma \cdot \mathbf{A} \cdot \mathbf{F}_\sigma^{-1}\right)(i, j) = \mathbf{A}(\sigma^{-1}(i), \sigma^{-1}(j))$.

Hint! If you work this correctly, and use the other exercises about permutation matrices, you only need to prove one of the three equalities.

Exercise 8.18. Show that the map

Solution

$$\begin{aligned} &: \mathbf{GL}_2(\mathbb{R}) \to \mathbf{PGL}_2(\mathbb{R}) \\ &: \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \frac{az + b}{cz + d} \end{aligned}$$

is an epimorphism with kernel equal to the set of scalar matrices: $\{r\mathbf{I}_2 : r \in \mathbb{R}^\times\}$. ($\mathbf{PGL}_2(\mathbb{R})$ was first defined in Exercise 5.3 back on page 240.)

(15) For any positive integer n, the function

$$\begin{aligned} \pi_n &: \mathbb{Z} \to \mathbb{Z}_n \\ &: m \mapsto [\, m \,]_n \end{aligned}$$

is a homomorphism of the additive groups. This is true by definition of the operation (addition) in $\mathbb{Z}_n$. In more detail, for every integer m and m' we have

$$\pi_n(m + m') = [\, m + m' \,]_n = [\, m \,]_n + [\, m' \,]_n = \pi_n(m) + \pi_n(m').$$

The kernel of this map is the set

$$\begin{aligned}\text{Ker}\,(\pi_n) &= \{m \in \mathbb{Z} : \pi_n(m) = [\,0\,]_n\}\\ &= \{m \in \mathbb{Z} : [\,m\,]_n = [\,0\,]_n\}\\ &= \{m \in \mathbb{Z} : n \mid m\}\\ &= n\mathbb{Z}.\end{aligned}$$

Not only does π_n preserve addition, it is also preserves (modular) multiplication. (It's not a homomorphism since $(\mathbb{Z}, \cdot)$ and $(\mathbb{Z}_n, \cdot)$ are not groups.)

(16) We generalize the previous example, **(15)**, a bit here. Let G be any group and let $g \in G$. We show that there are $|\,G\,|$ homomorphisms from $\mathbb{Z}$ to G. For instance there are 8 homomorphisms from $\mathbb{Z} \to \mathcal{D}_4$. First note that the function

$$\begin{aligned}\mathbf{E}_g &: \mathbb{Z} \to G\\ &: n \mapsto g^n.\end{aligned}$$

is a group homomorphism. This is because

$$\begin{aligned}\mathbf{E}_g(n+m) &= g^{n+m}\\ &= g^n \cdot g^m\\ &= \mathbf{E}_g(n) \cdot \mathbf{E}_g(m).\end{aligned}$$

The image of this homomorphism is the very definition of $\langle\, g\,\rangle$. (Note that as the image of some homomorphism, this gives us another proof that $\langle\, g\,\rangle$ is a group for every $g \in G$.) Moreover the kernel of this homomorphism is

$$\{n \in \mathbb{Z} : \mathbf{E}_g(n) = e_G\} = \{n \in \mathbb{Z} : g^n = e_G\}.$$

But this latter set is $|\,g\,|\mathbb{Z}$, the integer multiples of the order $|\,g\,|$, since $g^n = e_G$ happens iff the order of g, $|\,g\,|$, divides n. Note that this indeed does fit with the previous example since there the order of $[\,1\,]_n$ is n and we showed that the kernel was $n\mathbb{Z}$.

On the other hand, if φ is any group homomorphism from $\mathbb{Z}$ to G, we know that $\varphi(1)$ is defined and is equal to some $g \in G$, i.e. $\varphi(1) = g$. Then we must have that for any arbitrary integer n

$$\varphi(n) = (\varphi(1))^n = (g)^n = g^n.$$

where we have used the fact that φ is a homomorphism from the additive group $\mathbb{Z}$ to the multiplicative group G. But we also have that $\mathbf{E}_g(n) = g^n$ as well. Since n was an arbitrary integer, we know the equality of two functions: $\mathbf{E}_g = \varphi$. So this fact catches your attention I will write it in the following incredibly sexy way:

$$\begin{gathered}\text{Elements of } G \overset{\text{1-1}}{\leftrightsquigarrow} \hom(\mathbb{Z}, G).\\ g \mapsto \left(\mathbf{E}_g : n \mapsto g^n\right)\\ \varphi(1) \hookleftarrow \varphi.\end{gathered}$$

(17) The most important abstract isomorphism (i.e. one that can be defined for an arbitrary abstract group) is conjugation. It is actually an automorphism. To define this we fix an element g in a group G. Then

$$\begin{aligned}c_g &: G \to G &(1)\\ &: a \mapsto g\,a\,g^{-1}. &(2)\end{aligned}$$

- Now c_g is operation preserving since for all $a, b \in G$ we have (using $\cdot$ and parenthesis for emphasis)

$$\begin{aligned} c_g(ab) &= g\,(a \cdot b)\,g^{-1} \\ &= g\,(a \cdot e_G \cdot b)\,g^{-1} \\ &= g\left(a \cdot (g^{-1} \cdot g) \cdot b\right)g^{-1} \\ &= \left(g\,a\,g^{-1}\right) \cdot \left(g\,b\,g^{-1}\right) \qquad \text{(Re-associating)} \\ &= c_g(a)\,c_g(b).^{\text{II}} \end{aligned}$$

  Therefore c_g is operation preserving.

- We prove that c_g is a bijection by showing that it is invertible, with inverse $c_{g^{-1}}$. Indeed for all $a \in G$, we have that

$$\begin{aligned} \left(c_g \circ c_{g^{-1}}\right)(a) &= c_g(c_{g^{-1}}(a)) \\ &= c_g(g^{-1}ag) \\ &= g\left(g^{-1}ag\right)g^{-1} \\ &= \left(\cancel{gg^{-1}}\right)\ a\ \left(\cancel{gg^{-1}}\right) \\ &= a. \end{aligned}$$

  Therefore $c_g \circ c_{g^{-1}} = \mathbf{1}_G$. Since this equation holds for an arbitrary element of G it holds in particular for g^{-1}. Therefore

$$c_{g^{-1}} \circ c_{g^{-1}{}^{-1}} = \mathbf{1}_G \Rightarrow c_{g^{-1}} \circ c_g = \mathbf{1}_G$$

  where we have used the fact that $\left(g^{-1}\right)^{-1}$. Hence we have that $c_g^{-1} = c_{g^{-1}}$ showing that c_g is a bijection.

(18) Corollary 3.99 back on page 122 shows that

$$\|\ \|: \mathbb{H}^\times \to \mathbb{R}^+ \qquad \text{and its restriction} \qquad |\ |: \mathbb{C}^\times \to \mathbb{R}^+$$

are group homomorphisms. ($\mathbb{R}^+$ is the multiplicative group of positive real numbers).

$$\begin{aligned} \pi &: \mathbb{H}^\times \to \mathbb{S}^3 \\ &: q \mapsto \frac{q}{\|q\|} \end{aligned}$$

is also a homomorphism. Indeed, we have

$$\begin{aligned} \pi(q * q') &= \frac{q * q'}{\|q * q'\|} \\ &= \frac{q * q'}{\|q\|\,\|q'\|} \\ &= \frac{q}{\|q\|} * \frac{q'}{\|q'\|} \\ &= \pi(q) * \pi(q'). \end{aligned}$$

[II] This proof should be thought of as "inserting $g^{-1} \cdot g$ in the middle!

The last two examples are, in a sense, complimentary since

$$\begin{aligned}\operatorname{Ker}(\|\quad\|) &= \{\mathfrak{q} \in \mathbb{H}^\times : \|q\| = e_{\mathbb{R}^\times} = 1\} = \mathbb{S}^3 \\ \operatorname{Im}(\|\quad\|) &= \mathbb{R}^{+\,\text{III}}\end{aligned}$$

while

$$\begin{aligned}\operatorname{Ker}(\pi) = \left\{\mathfrak{q} \in \mathbb{H}^\times : \frac{\mathfrak{q}}{\|\mathfrak{q}\|} = e_{\mathbb{S}^3} = 1\right\} = \left\{\mathfrak{q} \in \mathbb{H}^\times : \mathfrak{q} = \|\mathfrak{q}\|\right\} = \mathbb{R}^+ \\ \operatorname{Im}(\pi) = \mathbb{S}^3.\end{aligned}$$

(19) In Exercise 2.64 way back on page 67, you showed that if $m \mid n$, that the map

$$\begin{aligned}\pi = \pi_n^m &: \mathbb{Z}_n \to \mathbb{Z}_m \\ &: [\,a\,]_n \mapsto [\,a\,]_m\end{aligned}$$

is well defined (and hence a function). To see that π is a homomorphism is now trivial, the proof is the same as showing that the function π was a homomorphism. It's kernel consists of

$$\begin{aligned}\{[\,a\,]_n \in \mathbb{Z}_n : \pi([\,a\,]_n) = [\,0\,]_m\} &= \{[\,a\,]_n \in \mathbb{Z}_n : [\,a\,]_m = [\,0\,]_m\} \\ &= \{[\,a\,]_n \in \mathbb{Z}_n : m \mid a\} \\ &= \{[\,0\,]_n, [\,m\,]_n, [\,2m\,]_n, \ldots\} \\ &= m\mathbb{Z}_n.\end{aligned}$$

Also note that not only does π preserve addition, it also preserves multiplication.

(20) Just like the previous example, suppose that $m \mid n$. Then the function

$$\begin{aligned}\Pi_n^m = \Pi &: \mathbf{SL}_2(\mathbb{Z}_n) \to \mathbf{SL}_2(\mathbb{Z}_n) \\ &: \begin{pmatrix}[\,a\,]_n & [\,b\,]_n \\ [\,c\,]_n & [\,d\,]_n\end{pmatrix} \mapsto \begin{pmatrix}[\,a\,]_m & [\,b\,]_m \\ [\,c\,]_m & [\,d\,]_m\end{pmatrix}\end{aligned}$$

is also a group homomorphism. (In other words it sends the matrix with coefficients $A_{i,j}$ to the matrix with coefficients $\pi_n^m(A_{i,j})$ defined in the previous example. That

[III] $\mathbb{R}^+$ denotes the positive real numbers, which incidentally form a group under multiplication.

is $\Pi(\mathbf{A})(i,j) = \pi(A_{ij})$. And since for any two matrices $\mathbf{A}$ and $\mathbf{B}$ in $\mathbf{SL}_2(\mathbb{Z}_n)$ we have that

$$\begin{aligned}\Pi(\mathbf{A}\cdot\mathbf{B})(i,k) = \pi((\mathbf{A}\cdot\mathbf{B})(i,k)) &= \pi\left(\sum_{j=1}^{2} A_{ij}B_{jk}\right)\\ &= \sum_{j=1}^{2}\pi\left(A_{ij}B_{jk}\right)\\ &= \sum_{j=1}^{2}\pi(A_{ij})\pi(B_{jk})\\ &= \sum_{j=1}^{2}\Pi(\mathbf{A})(i,j)\ \cdot\ \Pi(\mathbf{B})(j,k)\\ &= (\Pi(\mathbf{A})\cdot\Pi(\mathbf{B}))\,(j,k).\end{aligned}$$

Therefore, since all coefficients are the same, $\Pi(\mathbf{A}\cdot\mathbf{B}) = \Pi(\mathbf{A})\cdot\Pi(\mathbf{B})$. (The informal reason why this worked is because matrix multiplication consists of nothing besides multiplication and addition and the function π preserves both of these.)

EXERCISE 8.19. Explicitly write

$$c_{(1\,2\,3)}\colon \mathcal{S}_3 \to \mathcal{S}_3$$

down by writing the image (under $c_{(1\,2\,3)}$) of every element of $\mathcal{S}_3$ down. For example

$$(1\,2\,3)(1\,2)(1\,2\,3)^{-1} = (2\,3) \qquad \Rightarrow \qquad c_{(1\,2\,3)}\colon (1\,2) \mapsto (2\,3).$$

SOLUTION

EXERCISE 8.20.

(1) Show that the map

$$\begin{aligned}\hat{}^2\colon \mathcal{D}_4 &\to \mathcal{D}_4\\ \colon w &\mapsto w^2\end{aligned}$$

is not a homomorphism.

(2) Show that the map

$$\begin{aligned}\hat{}^3\colon \mathcal{D}_4 &\to \mathcal{D}_4\\ \colon w &\mapsto w^3\end{aligned}$$

is not a homomorphism.

SOLUTION

(3) Show that the map

$$\begin{aligned} \hat{}^5 &: \mathcal{D}_4 \to \mathcal{D}_4 \\ &: w \mapsto w^3 \end{aligned}$$

is a homomorphism.

Hint! Don't waste too much time! All these maps coincide with previously discussed maps

THEOREM 8.22. *Suppose that*

$$\varphi : G \to H$$

is a homomorphism between groups G and H. Then

(1) *If $h = \varphi(g)$, then $|h|$ divides $|g|$.*

(2) *If moreover φ is an isomorphism, then $|h| = |g|$.*

PROOF.

(1) Let's denote $|g|$ as n. Then we have that

$$e_H = \varphi(e_G) = \varphi(g^n) = \varphi(g)^n = h^n$$

where we have used Theorem 8.10**(3)** to get the third equality. But this implies that $n \Big| |h|$ by Theorem 5.38**(2)**, which is what we wanted to show.

(2) Suppose that φ is an isomorphism. We already know that $|h| \Big| |g|$ by the first part. You've shown in Excercise 8.10**(??)** that φ^{-1} is also an isomorphism. For emphasis, we let ψ denote this inverse, φ^{-1}. Since $\psi(h) = \psi(\varphi(g)) = g$, we can again apply the first part of this theorem to conclude that $|g| \Big| |h|$ as well. Since both numbers are positive and divide each other, they are equal. ∎

REMARK 8.23. For an alternative proof of the second part, note that if $\varphi : G \to H$ is an isomorphism, then we know

In symbols, the following corollary says

$$G \cong \widehat{G} \Rightarrow \forall n \in \mathbb{Z}^+, \mathbf{o}_n(G) = \mathbf{o}_n(\widehat{G}).$$

COROLLARY 8.24. *Suppose we have two isomorphic groups G and $\widehat{G}$. Let n be an integer. Then the number of elements of order n in G is the same as the number of elements of order n in $\widehat{G}$.*

PROOF. The theorem just presented, Theorem 8.22, says that an element $g \in G$ has order n iff its image $\varphi(g)$ has order n. Thus, φ being a bijection guarantees the result. ∎

COROLLARY 8.25. *Suppose $\varphi : G \to H$ is an epimorphism (and not necessarily a bijection) of finite groups. Suppose that H has an element of order n, then so does G.*

PROOF. Suppose that H contains an element of order n. Let $h \in H$ be such an element. Then, since we are assuming that φ is surjective, there exists an element $g \in G$ so that $\varphi(g) = h$. By the theorem, we have that the order of h, which is n, divides the order of g. Then, as we saw in Lemma 5.77 back on page 291, $g^{\frac{|g|}{n}}$ is an element of G of order n. ∎

Example 8.26. The only homomorphism $\mathcal{S}_3 \to \mathbb{Z}_5$ is the trivial homomorphism, i.e. the homomorphism which sends every element to the identity (of $\mathbb{Z}_5$). This is because every element of of $\mathbb{Z}_5$, except for the identity, has (additive!) order 5. However, no element of $\mathcal{S}_3$ has order 5. Therefore the image of any homomorphism $\mathcal{S}_3 \to \mathbb{Z}_5$ must be contained inside of $\{[\,0\,]_5\}$, i.e. any homomorphism must be trivial.

Exercise 8.21. Suppose that φ is an epimorphism of groups $G \to H$. Prove that

$$z \in \mathcal{C}(G) \Rightarrow \varphi(z) \in \mathcal{C}(H).$$

(So you are proving that $\varphi(\mathcal{C}(G)) \subseteq \mathcal{C}(H)$.)

Solution

Exercise 8.22. Show that the conclusion of the previous exercise fails if we do not assume that φ is surjective.
Hint! Pick any abelian subgroup of a centerless group and consider the inclusion.

Solution

Exercise 8.23. Find an example of an epimorphism $G \to H$ with some $h \in \mathcal{C}(H)$ with no $g \in \mathcal{C}(G)$ mapping to h.
Clarification! In other words show that $\varphi|_{\mathcal{C}(G)} \colon \mathcal{C}(G) \to \mathcal{C}(H)$ is not necessarily surjective even when the map $\varphi \colon G \to H$ is.

Solution

Theorem 8.27. *If $G \cong \dot{G}$, then $\mathcal{C}(G) \cong \mathcal{C}(\dot{G})$.*

Proof. By the previous exercise, Exercise 8.21, we know that an isomorphism (being an epimorphism) $\varphi \colon G \to \dot{G}$ takes the center of G to the center of $\dot{G}$, and so it restricts to a monomorphism

$$\varphi|_{\mathcal{C}(G)} \colon \mathcal{C}(G) \to \mathcal{C}(\dot{G}).$$

We simply need to show that this restriction is surjective. To that end, let $z \in \mathcal{C}(\dot{G})$ be arbitrary. By a different problem, Exercise 8.10**(1)**, we know that φ^{-1} is an isomorphism $\dot{G} \to G$. Therefore, using Exercise 8.21 again (this time applied to φ^{-1} instead of φ) implies that $z' := \varphi^{-1}(z) \in \mathcal{C}(G)$. But $\varphi(z') = \varphi(\varphi^{-1}(z)) = z$. Since z was arbitrary, $\varphi|_{\mathcal{C}(G)}$ is surjective, completing the proof that it is an isomorphism,

$$\varphi|_{\mathcal{C}(G)} \colon \mathcal{C}(G) \cong \mathcal{C}(\dot{G}).$$

■

In particular, this means that there is a bijection[IV] between the centers of two isomorphic groups, hence such groups have the same number of elements in their centers. Therefore we have proven:

Corollary 8.28. *The number of elements in the centers of isomorphic groups are the same.* ■

This is often the fastest way of proving that two groups are not isomorphic. For instance, the group $\mathcal{D}_6$ and $\mathcal{A}_4$ each have order 12. However $\mathcal{D}_6$ has a center containing two elements (the non–identity element being rotation by 180°), while $\mathcal{A}_4$ has a center containing only 1 (it is centerless). Thus $\mathcal{D}_6$ and $\mathcal{A}_4$ are non-isomorphic groups of order 12.

[IV] An isomorphism is a special kind of bijection.

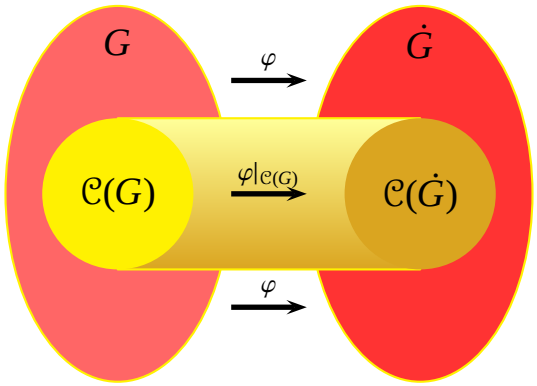


Figure 8.1: $G \cong \dot{G} \Rightarrow \mathcal{C}(G) \cong \mathcal{C}(\dot{G})$

The following Example/Theorem is a bit more abstract than other examples we have encountered. It deals with a map between any group G to its group of automorphisms, $\mathrm{Aut}(G)$, defined just above in Exercise 8.10.

Recall that the center of a group G is the set of elements in G which commute with every element of G.

EXAMPLE /THEOREM 8.29. The map $g \mapsto c_g$ defines a homomorphism $G \to \mathrm{Aut}(G)$ with kernel equal to the center of G, $\mathcal{C}(G)$.

PROOF. By **(17)** on page 391 we know that for any $g \in G$, that c_g is an isomorphism $G \to G$. This means, by definition, that $c_g \in \mathrm{Aut}(G)$ and the function $g \mapsto c_g$ has image contained in $\mathrm{Aut}(G)$. We need to show that it is operation preserving (and hence a homomorphism) and then calculate its kernel.

- Let $a, b \in G$ be arbitrary. Then for all $g \in G$ we have that

$$\begin{aligned} c_{ab}(g) &= (ab)\, g\, (ab)^{-1} \\ &= ab\, g\, b^{-1}a^{-1} \qquad \text{(Socks and Shoes)} \\ &= a\,(bgb^{-1})\, a^{-1} \\ &= c_a(bgb^{-1}) \\ &= c_a(c_b(g)) \\ &= (c_a \circ c_b)\,(g). \end{aligned}$$

  Since g was arbitrary, we have proven that $c_{ab} = c_a \circ c_b$. Since $\circ$ is the operation in the group $\mathrm{Aut}(G)$, $g \mapsto c_g$ is operation preserving.

- The group $\mathrm{Aut}(G)$ is a group of functions which has an identity equal to the identity function, $\mathbf{1}_G$. Therefore, the kernel of the homomorphism $g \mapsto c_g$ is the set of

$g \in G$ so that $c_g = \mathbf{1}_G$. So

$$\begin{aligned}\operatorname{Ker}\left(g \mapsto c_g\right) &= \{g \in G : c_g = \mathbf{1}_G\} \\ &= \{g \in G : \forall a \in G, c_g(a) = a\} \\ &= \{g \in G : \forall a \in G, gag^{-1} = a\} \\ &= \{g \in G : \forall a \in G, ga = ag\} \\ &= \mathcal{C}(G).\end{aligned}$$

■

EXAMPLE 8.30. Let **F** be the homomorphism from **(14)** on page 388 which takes a permutation to its permutation matrix. There is a homomorphism given by the composition (the composition of two homomorphisms is a homomorphism by

$$\mathcal{S}_n \xrightarrow{\mathbf{F}} \mathbf{GL}_n(\mathbb{Z}) \xrightarrow{\det} \mathcal{U}(\mathbb{Z}) = \{\pm 1\}.$$

This homomorphism coincides with the ε function defined way back in chapter 1 which simply counted the parity of the number of crossings in the braid diagram of a permutation. Since we weren't very concise when defining the determinant, and since we will use ε for theoretical, and not just example based, part of the text we stick with this and only the first, inductive, definition of the determinant.

The kernel of the sign homomorphism $\mathcal{S}_n \to \{\pm 1\}$, is by the very definition given the set of even permutations. This is because the identity of the group $\pm 1 = \{+1, -1\}$ is the number $+1 = 1$. Recall that It is denoted by $\mathcal{A}_n$ and is called *the alternating group*.

DEFINITION 8.31. The homomorphism defined in the previous example, 8.30 is denoted by ε. This replaces the older definition for the time being. We carry over all the definition, including parity of a permutation, even permutations, and odd permutations exactly as they were before. .

EXERCISE 8.24. (Direct Product Homomorphisms)

SOLUTION

(1) Prove that $H \times K \cong K \times H$.

(2) More generally, prove that for any collection of groups $H_1, \ldots, H_n$ and permutation $\sigma \in \mathcal{S}_n$ the groups

$$H_1 \times H_2 \times \cdots \times H_n \cong H_{\sigma(1)} \times H_{\sigma(2)} \times \cdots \times H_{\sigma(n)}.$$

(3) Prove that

$$H \cong H', K \cong K' \Rightarrow H \times K \cong H' \times K'.$$

(4) Prove that

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

(5) Suppose that

$$\varphi\colon G \to H \quad \text{and} \quad \psi\colon G \to K$$

are group homomorphisms. Show that

$$\begin{aligned}(\varphi,\psi)\colon G &\to H\times K\\ g &\mapsto (\varphi(g),\psi(g))\end{aligned}$$

is a homomorphism with

$$\operatorname{Ker}((\varphi,\psi)) = \operatorname{Ker}(\varphi)\cap\operatorname{Ker}(\psi).$$

(6) Show each of the maps

$$H \xrightarrow{i_H} H\times K \xrightarrow{\pi_K} K$$

defined by $i_H(h) = (h, e_K)$ and $\pi_K(h,k) = k$ are homomorphisms. Show that

$$H \cong \operatorname{Im}(i_H) = \operatorname{Ker}(\pi_K).$$

EXERCISE 8.25. The following statement (and hence its proof) is false. Find the error in the proof.

FALSE THEOREM 8.32. *For any integers n and m*

$$\begin{aligned}\pi\colon \mathbb{Z}_n &\to \mathbb{Z}_m\\ :\ [\,a\,]_n &\mapsto [\,a\,]_m\end{aligned}$$

is a group homomorphism.

"**PROOF**". For any $[\,a\,]_n$ and $[\,a'\,]_n$ in $\mathbb{Z}_n$ we have that

$$\begin{aligned}\pi([\,a\,]_n + [\,a'\,]_n) &= \pi([\,a+a'\,]_n) \qquad (3)\\ &= [\,a+a'\,]_m = [\,a\,]_m + [\,a\,]_m\\ &= \pi([\,a\,]_n) + \pi([\,a'\,]_n).\end{aligned}$$

Therefore, π is a group homomorphism as desired. ■

Hint! There is nothing wrong with the calculation done in (3).

SOLUTION

The following theorem is a very important theorem in these notes. It should also be far more intuitive than its proof. It says that if we have n things, it doesn't matter what we call those n things when we permute them. There could be n dogs, n rocket ships, or n math dorks, but exchanging dog 1 with dog 2 is the same as exchanging rocket ship 1 with rocket ship 2 which is the same as exchanging 1 math dork for another.

THEOREM 8.33. *If X is a finite set with n-elements, then $\mathcal{S}_X$ is isomorphic to $\mathcal{S}_n$.*

Proof. [V] Label the n elements of X by the numbers 1 through n in any way you would like. More technically, we pick a dictionary (i.e. a bijection)

$$F : \overline{n} \to X$$
$$F : i \mapsto x_i.$$

We wish to define an isomorphism Φ from $\mathcal{S}_n$ to $\mathcal{S}_X$. If we have a permutation $\sigma \in \mathcal{S}_n$ which sends the letter i to the letter j, $\Phi(\sigma)$ is defined to be the permutation which takes x_i to x_j. To make the calculations which follow easier, we derive a formula for Φ.

$$\begin{array}{ccc} \overline{n} & \xrightarrow{F} & X \\ {\scriptstyle\sigma}\downarrow & & \downarrow{\scriptstyle\Phi(\sigma)} \\ \overline{n} & \xrightarrow{F} & X \end{array} \tag{4}$$

I claim that this diagram is commutative (which is a visual way of saying that $\Phi(\sigma) \circ F = F \circ \sigma$) since if we feed the left hand side i, we get

$$\Phi(\sigma)(F(i)) = \Phi(\sigma)(x_i) = x_j$$

and if we feed the right hand side this same i we also get

$$F(\sigma(i)) = F(j) = x_j.$$

Since functions are determined by what they do to elements of their domain (the domain of both $\Phi(\sigma) \circ F$ and $F \circ \sigma$ is $\overline{n}$) the diagram, (4) is indeed commutative. After composing both sides with F^{-1} (which exists since F is a bijection) we therefore get the formula for $\Phi(\sigma)$ we were after:

$$\Phi(\sigma) = F \circ \sigma \circ F^{-1}.$$ [VI] (5)

For one thing, this formula says that F is a permutation (of X) (i.e. a bijection $X \to X$) whenever σ is (a permutation of $\overline{n}$) because it shows that $F(\sigma)$ is a composition of the bijections – F, σ, and F^{-1} – and is hence a bijection. Therefore we know for certain that Φ is a function

$$\Phi : \mathcal{S}_n \to \mathcal{S}_X.$$

The rest of the proof that Φ is an isomorphism is exactly the same as proving that conjugation by g was an isomorphism back on page 391. [VII] We just illustrate this by showing that Φ is operation preserving, and leave the fact that it is a bijection to you (again, its simply a translation of the previous proof).

- Φ preserves operations since given elements $\sigma, \tau \in \mathcal{S}_n$ we have

$$\begin{aligned} \Phi(\sigma \circ \tau) &= F \circ (\sigma \circ \tau) \circ F^{-1} && \text{(By (5))} \\ &= F \circ (\sigma \circ \left(F^{-1} \circ F\right) \circ \tau) \circ F^{-1} \\ &= \left(F \circ \sigma \circ F^{-1}\right) \bullet \left(F \circ \tau \circ F^{-1}\right) \\ &= \Phi(\sigma) \circ \Phi(\tau). \end{aligned}$$

Therefore we have proven that Φ is operation preserving as desired. ∎

[V] We present this proof in a less formal way than usual.
[VI] Of course I could have done a proof by intimidation and simply defined $\Phi(\sigma)$ by this formula, but students typically understand the pseudo derivation we have just done better
[VII] It would be a special case of it if the set of all bijections (between arbitrary sets) were a group, but alas, it is not ... the difference being that we cannot compose bijections unless the domains line up. Such things are called *groupoids* in the literature.

Remark 8.34.

(1) Unfortunately, this proof makes this fact seem much more complicated than it really is. As pointed out in the proof, in practice, you end up doing what you'd expect you would do. For instance, the isomorphism would send a permutation of $\overline{n}$ written in cycle notation to the permutation of X with cycle notation which simply replaces i with x_i. For instance

$$(1\ 2\ 5)(3\ 7) \mapsto (x_1\ x_2\ x_5)(x_3\ x_7).$$

(2) Note that the isomorphism $\mathcal{S}_n \to \mathcal{S}_X$ depends on the choice of labeling $F\colon X \to \overline{n}$.[VIII] As in Chapter 1, we referred to such a labeling as a *dictionary.* Note that in the special case that X also equals $\overline{n}$, a bijection $\sigma\colon \overline{n} \to \overline{n}$ induces an isomorphism $\mathcal{S}_n \to \mathcal{S}_n$ which is given by conjugation by σ.

Exercise 8.26. Explicitly calculate its kernel of the homomorphism

$$\Pi\colon \mathbf{SL}_2(\mathbb{Z}_4) \to \mathbf{SL}_2(\mathbb{Z}_2)$$

defined in the examples on page 393 just above.
Hint! There are a total of 8 elements in its kernel. The matrix

$$\begin{pmatrix} [3]_4 & [2]_4 \\ [0]_4 & [3]_4 \end{pmatrix}$$

is an example of an element of the kernel since $[3]_2 = [1]_2$ and $[2]_2 = [0]_2$.

SOLUTION

Exercise 8.27. Define

$$\begin{aligned} \varphi\colon \mathbf{SL}_2(\mathbb{Z}_6) &\to \mathbf{SL}_2(\mathbb{Z}_2) \times \mathbf{SL}_2(\mathbb{Z}_3) \\ : \begin{pmatrix} [a]_6 & [b]_6 \\ [c]_6 & [d]_6 \end{pmatrix} &\mapsto \left(\begin{pmatrix} [a]_2 & [b]_2 \\ [c]_2 & [d]_2 \end{pmatrix}, \begin{pmatrix} [a]_3 & [b]_3 \\ [c]_3 & [d]_3 \end{pmatrix} \right). \end{aligned}$$

(1) Prove that φ is a group homomorphism.

(2) Prove that φ is injective. [IX]

SOLUTION

Let's collect all of these facts and express them in terms of the conjugation automorphisms c_g. Notice that many of these facts are things I had you prove independently in Exercise 5.61.

Theorem 8.35. *For any $g, a \in G$ and $H \leqslant G$ we have*

(1) $ga^{-1}g^{-1} = \left(gag^{-1}\right)^{-1}$

(2) $\left|gag^{-1}\right| = |a|$

[IX] Later we will show that this map is also surjective which will prove that φ defines an isomorphism.

(3) $g\,a^n\,g^{-1} = \left(g\,a\,g^{-1}\right)^n$

(4) *The set*

$$gHg^{-1} \overset{def}{=} \{ghg^{-1} : h \in H\}$$

is also a subgroup of G. Moreover, it is isomorphic to H.

Proof. Part **(4)** follows directly from the fact that

$$gHg^{-1} = c_g(H)$$

and as such, it is the image of a subgroup (namely H) under a monomorphism (namely c_g). Other parts are direct translations of the various theorems, lemmas, and propositions proven in this chapter but applied to the specific isomorphism c_g. ■

It's interesting to try and prove these directly. If you spend a moment doing this (correctly!) you should quickly realize that you are following the proof of 8.10 pretty closely.

Example 8.36.

(1) We just proved that if A and $B \subseteq \overline{n}$ have the same number of elements, then $\mathcal{S}_A$ and $\mathcal{S}_B$ are conjugate subgroups of $\mathcal{S}_n$. In fact we showed that if $\sigma \in \mathcal{S}_n$ satisfies $\sigma(A) = B$, then $\sigma\mathcal{S}_A\sigma^{-1} = \mathcal{S}_B$.

(2) A similarity of $\mathbb{R}^3$ is a function $\alpha\colon \mathbb{R}^3 \to \mathbb{R}^3$ with the property that there exists an $r \in \mathbb{R}^+$ so that for all $\mathbf{v} \in \mathbb{R}^3$, $\|\alpha(\mathbf{v})\| = r \cdot \|\mathbf{v}\|$. Isometries are similarities with $r = 1$. The function $\mathbf{v} \mapsto 2 \cdot \mathbf{v}$ is a similarity which is not an isometry. Then since we have for an isometry φ

$$\left\|\alpha \circ \varphi \circ \alpha^{-1}(\mathbf{v})\right\| = r \cdot \left\|\varphi \circ \alpha^{-1}(\mathbf{v})\right\| = r \cdot \left\|\alpha^{-1}(\mathbf{v})\right\| = r \cdot \frac{1}{r}\|\mathbf{v}\| = \|\mathbf{v}\|$$

we have that $\alpha \circ \varphi \circ \alpha^{-1}$ is an isometry. This implies that if two geometric shapes X and Y in $\mathbb{R}^3$ we have a similarity $\alpha\colon \mathbb{R}^3 \to \mathbb{R}^3$ with $\alpha(X) = Y$ (said that X and Y are similar), then their symmetry groups are conjugate:

$$\alpha\, S(X)\, \alpha^{-1} = S(Y) \quad \text{and} \quad \alpha\, S^*(X)\, \alpha^{-1} = S^*(Y).$$

(The proof is basically done, I leave the details to you.) This is the formal reason why it doesn't matter "which" regular triangle we study when studying the symmetries of "a triangle" since all triangles are similar. More generally, and two regular polygons of the same type are similar and the same goes for any to platonic solids of the same type.

Exercise 8.28. Let

$$B := \{(\mathsf{z}, \sigma) \in \{\{4\}\} \times \mathcal{S}_3 : \mathsf{z}^2 = \varepsilon(\sigma)\}.$$ [x]

SOLUTION

(1) Prove that B is a group.

(2) Define a function $\pi_1\colon B \to \{\{4\}\}$ which sends (z, σ) to z. Prove that this map is an epimorphism whose kernel contains 3 elements.

(3) Define a function $\pi_2\colon B \to \mathcal{S}_3$ which sends (z, σ) to σ. Prove that this map is an epimorphism whose kernel contains 2 elements.

(4) Prove that for every $\sigma \in \mathcal{S}_3$, $\#(\pi_2^{-1}(\sigma)) = 2$.

(5) Prove that $|B| = 12$.

Exercise 8.29. Fix a positive integer $n > 1$. Define a map

$$\begin{aligned}\varphi\colon \mathbf{GL}_{n-1}(R) &\to \mathbf{M}_n(R)\\ \colon \mathbf{A} &\mapsto \begin{pmatrix}\mathbf{A} & 0\\ 0 & \det \mathbf{A}^{-1}\end{pmatrix}.\end{aligned}$$ [XI]

(1) Prove φ defines a monomorphism

$$\mathbf{GL}_{n-1}(R) \to \mathbf{SL}_n(R).$$

(2) Conclude that $\mathbf{GL}_{n-1}(R)$ is isomorphic to a subgroup of $\mathbf{SL}_n(R)$.

SOLUTION

8.4 Normal and Malnormal Subgroups

Let H be a subgroup of a group G. Remember from Theorem 8.35**(4)** for any $g \in G$ we have that

$$gHg^{-1} \stackrel{\text{def}}{=} \{ghg^{-1} : h \in H\}$$

is also a subgroup of G which is isomorphic to the original subgroup, H. We generically call such subgroups *conjugate* subgroups of H. The intersection of all conjugate subgroups of H is called the *core* of H (in G). It is denoted as $H^{\circ G}$ or simply as H° (if G is understood). In symbols,

Definition 8.37.

$$H^{\circ G} \stackrel{\text{def}}{=} \bigcap_{g \in G} gHg^{-1}.$$

Note that since every element of a group has a unique inverse, the set $\{g : g \in G\}$ is the same as the set $\{g : g^{-1} \in G\}$. Indeed, both of course are equal to G. Therefore we could just have well taken

$$H^{\circ G} \stackrel{\text{def}}{=} \bigcap_{g \in G} g^{-1}Hg$$

[X] For example: $(i, (1\ 2))$, $(-1, (1\ 2\ 3))$, and $(1, (1\ 2\ 3)) \in B$–but $(i, (1\ 2\ 3))$ and $(1, (1\ 2))$ are not. We'll eventually see that $B \cong B\mathcal{D}_3$.

[XI] This is the matrix whose upper left hand $n-1$ by $n-1$ block coincides with $\mathbf{A}$, whose lower right hand entry is $\det \mathbf{A}^{-1}$ (which makes sense since we're assuming $\mathbf{A} \in \mathbf{GL}$ ()) and all other entries are 0. So for example, if $n = 3$ we have

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix} \mapsto \begin{pmatrix}a & b & 0\\ c & d & 0\\ 0 & 0 & (ad-bc)^{-1}\end{pmatrix}.$$

in Definition 8.37. This means that

$$\begin{aligned} h \in H^{\circ G} &\Leftrightarrow \forall g \in G,\ h \in g^{-1}Hg \\ &\Leftrightarrow \forall g \in G \exists \widehat{h} \in H,\ h = g^{-1}\widehat{h}g \\ &\Leftrightarrow \forall g \in G \exists \widehat{h} \in H,\ ghg^{-1} = \widehat{h} \\ &\Leftrightarrow \forall g \in G \exists \widehat{h} \in H,\ ghg^{-1} \in H. \end{aligned}$$

In other words

PROPOSITION 8.38.

$$H^{\circ G} = \{h \in H : \text{ for all } g \in G,\ ghg^{-1} \in H\}.$$

■

EXERCISE 8.30. Suppose that $H \leqslant G$, then

(1) $H^{\circ G} \leqslant H$

(2) $H^{\circ G} \leqslant G$.

SOLUTION

Hint! If you cite the proper theorems each part of this exercise'll only take a single line.

EXAMPLE 8.39. In this example we take

$$G = \mathcal{S}_4 \text{ and } H = \langle (1\,2\,3\,4) \rangle = \{(1), (1\,2\,3\,4), (1\,3)(2\,4), (1\,4\,3\,2)\}.$$

Then $(1\,2\,3\,4) \notin H^{\circ G}$ since

$$(1\,2\,3)(1\,2\,3\,4)(1\,2\,3)^{-1} = (2\,3\,1\,4) = (1\,4\,2\,3)$$

$$H^{\circ} \stackrel{?}{=} \{(1), \cancel{(1\,2\,3\,4)}, (1\,3)(2\,4), (1\,4\,3\,2)\}$$

is not a power of $(1\,2\,3\,4)$ and so not in H. ⟨Therefore there exists a $g \in G$ (namely $(1\,2\,3)$) so that $g(1\,2\,3\,4)g^{-1} \notin H$. ⟩

At the same time, $(1\,2\,3\,4)^{-1} = (1\,4\,3\,2)$ is also not in the core of H since we have that $(1\,2\,3)(1\,4\,3\,2)(1\,2\,3)^{-1} = (1\,3\,2\,4)$ is not a power of H. (Alternatively, since the core is a subgroup and we've already shown the inverse of $(1\,4\,3\,2)$ is not in the core, we already knew that $(1\,2\,3\,4)$ was not going to be in the core.)

$$H^{\circ} \stackrel{?}{=} \{(1), \cancel{(1\,2\,3\,4)}, (1\,3)(2\,4), \cancel{(1\,4\,3\,2)}\}$$

Lastly $(1\,3)(2\,4)$ is also not in the core since, [I]

$$(1\,2)(1\,3)(2\,4)(1\,2)^{-1} = (2\,3)(1\,4)$$

which is not in H.

[I] We're using a different g not because we have to, indeed $(1\,2\,3)$ will again do the job, but to avoid setting a precedent which might lead you to believe that the same $g \in G$ must be used for all $h \in H$.

$$H^\circ \stackrel{?}{=} \{(1), (1\,2\,3\,4), (1\,3)(2\,4), (1\,4\,3\,2)\} \Rightarrow H^\circ = \{(1)\}.$$

Therefore the core of H is trivial.

Example 8.40. Consider the subgroup

$$H := \langle \mathbf{j} \rangle = \{1, \mathbf{j}, -1, -\mathbf{j}\} \leqslant B\mathcal{D}_3.$$

Then we have (denoting $e^{\pi \mathbf{i}/3}$ as χ) we have

$$\begin{aligned} 1H1^{-1} &= \{1, \mathbf{j}, -1, -\mathbf{j}\} = H \\ \chi H \chi^{-1} &= \chi\{1, \mathbf{j}, -1, -\mathbf{j}\}\chi^{-1} = \{1, \chi\mathbf{j}\chi^{-1}, -1, -\chi\mathbf{j}\chi^{-1}\} = \{1, \chi^2\mathbf{j}, -1, -\chi^2\mathbf{j}\} \\ \chi^2 H \chi^{-2} &= \chi^2\{1, \mathbf{j}, -1, -\mathbf{j}\}\chi^{-2} = \{1, \chi^4\mathbf{j}, -1, -\chi^4\mathbf{j}\} = \{1, -\chi\mathbf{j}, -1, \chi\mathbf{j}\}. \end{aligned}$$

It is not difficult to check that these are the only possible conjugates. For example, since $\chi^5 = \chi^3\chi^2 = -\chi^2$ we have that

$$\chi^5 H \chi^{-5} = \chi^2(-1)H(-1)\chi^{-2} = \chi^2 H \chi^{-2}$$

which was already written down. More formally, this is a consequence of the fact that in $B\mathcal{D}_3$ every element can be written as $\chi^m \mathbf{i}^n$ with $0 \leq m < 3$ and $0 \leq n < 4$. Therefore, inspecting these conjugates, we see that the core is just

$$H^{\circ B\mathcal{D}_3} = \{1, -1\} = \{\pm 1\}.$$

Therefore, the largest the core of a subgroup can be is the subgroup itself and the smallest the core of a subgroup can be is (like any subgroup) just the identity. These two extreme cases are given special names.

Definition 8.41. A subgroup $H \leqslant G$ is

- *Normal* if $H^{\circ G} = H$
- *Malnormal* if instead $H^{\circ G} = \{e_G\}$. [II]

For instance, in an abelian group G we have all subgroups are normal. This is because in such a situation, we have for all $g \in G$,

$$\begin{aligned} gHg^{-1} &= \{ghg^{-1} : h \in H\} \\ &= \{gg^{-1}h : h \in H\} \qquad \text{(Since } G \text{ assumed abelian)} \\ &= \{h : h \in H\} \\ &= H. \end{aligned}$$

Therefore, each of the gHg^{-1}'s are exactly the same set, namely H itself, and

$$H^{\circ G} = \bigcap_{g \in G} gHg^{-1} = \bigcap_{g \in G} H = H.$$

We record this as a proposition:

[II] In the literature, what we are calling a malnormal subgroup others often call, (less imaginatively but more descriptively!) a core-less subgroup. They reserve the "malnormal" title for subgroups which satisfy the stronger condition that

$$gHg^{-1} \cap H$$

is non-trivial iff $g \in H$. Or terminology hopefully highlights the dichotomy between malnormal and normal subgroups.

PROPOSITION 8.42. *Every subgroup of an abelian group is normal.* ■

To show that a subgroup H is malnormal we of course do not need to compute all conjugates! As soon as we find *some* conjugates of H which intersection is trivial (i.e. contains only e_G) we can conclude that H is malnormal. Often we can get away with just calculating just a single conjugate! Therefore a method for proving that H is malnormal is to find some element $g \in G$ so that gHg^{-1} and the original group H have only the identity in common. Of course, this might not always work, but for many examples in this book it does indeed work.

WORKING DEFINITION OF $H \trianglelefteq G$

PROPOSITION 8.43. *A subgroup H of a group G is normal iff*

$$\forall h \in H, \forall g \in G \; ghg^{-1} \in H.$$

PROOF. This is just a restatement of Proposition 8.38. ■

EXAMPLE 8.44. I show that $H = \langle (1\,2\,3\,4) \rangle$ is a malnormal subgroup of $\mathcal{S}_4$. Indeed, this is just a restatement of Example 8.39. If you want to see this all done at once note that

$$\begin{aligned}
\sigma H \sigma^{-1} &= \sigma \{(1\,2\,3\,4), (1\,3)(2\,4), (1\,4\,3\,2), (1)\} \sigma^{-1} \\
&= \{\sigma(1\,2\,3\,4)\sigma^{-1}, \sigma(1\,3)(2\,4)\sigma^{-1}, \sigma(1\,4\,3\,2)\sigma^{-1}, \sigma(1)\sigma^{-1}\} \\
&= \{(1\,2)(1\,2\,3\,4)(1\,2)^{-1}, (1\,2)(1\,3)(2\,4)(1\,2)^{-1}, (1\,2)(1\,4\,3\,2)(1\,2)^{-1}, (1\,2)(1)(1\,2)^{-1}\} \\
&= \{(2\,1\,3\,4), (2\,3)(1\,4), (2\,4\,3\,1), (1)\} \\
&= \{(1\,3\,4\,2), (1\,4)(2\,3), (1\,2\,4\,3), (1)\}.
\end{aligned}$$

SOLUTION

EXERCISE 8.31. As usual, we think of $\mathcal{S}_{n-1}$ as the subgroup of $\mathcal{S}_n$ consisting of permutations which send the letter n to itself. We have seen that for a permutations τ and for an element $\sigma \in \mathcal{S}_{n-1}$ that $\tau\sigma\tau^{-1}$ now fixes $\tau(n)$. This is because

$$\tau\sigma\tau^{-1}(\tau(n)) = \tau\sigma(n) = \tau(n).$$

Here we used the assumption that $\sigma(n) = n$. Therefore the subgroup $\tau\mathcal{S}_{n-1}\tau^{-1}$ is the subgroup of $\mathcal{S}_n$ which fixes $\tau(n)$. Since we can pick a τ which takes n to any letter in $\overline{n}$ we'd like we would have

$$\bigcap_{\tau \in \mathcal{S}_n} \tau\mathcal{S}_{n-1}\tau^{-1} = \bigcap_{\tau \in \mathcal{S}_n} \text{Perms in } \mathcal{S}_n \text{ fixing } \tau(n) = \text{Perms fixing every } i \in \overline{n}.$$

Of course only the identity fixes every letter in $\overline{n}$ and therefore this intersection is trivial. This is exactly what we needed to conclude that $\mathcal{S}_{n-1} \leqslant \mathcal{S}_n$ is malnormal.

SOLUTION

EXERCISE 8.32.

(1) In Example 10.74 coming up on page 535, we will see that every subgroup of Q_8 is normal[III]. Verify this by showing that the following subgroups of Q_8 are normal:

- $\langle \mathbf{i} \rangle$
- $\langle \mathbf{j} \rangle$
- $\langle -1 \rangle$

(2) We will see that every proper subgroup of $\mathcal{A}_4$ is malnormal. Verify this by showing that $\langle (1\,2\,3\,4\,5) \rangle$ is a malnormal subgroup of $\mathcal{A}_4$. (Be sure to show that whatever elements you use to conjugate this subgroup really are even permutations!)

The definition of normal requires that $H \subseteq gHg^{-1}$ for all $g \in G$. The following proposition gives us some leeway with this. For instance, this proposition is really obvious when H is finite since in this case H and gHg^{-1} have the same number of elements, so if one is contained inside of the other, the two sets must be equal.

PROPOSITION 8.45. TFAE

(1) *For all $g \in G$, $gHg^{-1} \subseteq H$*

(2) *For all $g \in G$, $H \subseteq gHg^{-1}$*

(3) *For all $g \in G$, $gHg^{-1} = H$.*

PROOF. Part **(3)** is really parts **(1)** and **(2)** written together. Therefore, we only need to show that **(1)** and **(2)** are equivalent.

1 ⇒ 2 Let $h \in H$ and $g \in G$ be arbitrary. We need to find an $h' \in H$ so that $h = gh'g^{-1}$. But since we are assuming that **(1)** holds, and since $g^{-1} \in G$ we know that $g^{-1}Hg \subseteq H$. This means that there exists an $h' \in H$ with $g^{-1}hg = h'$, but this proves that $h = gh'g^{-1}$ as desired.

2 ⇒ 1 The proof of this part is very similar. Let $ghg^{-1} \in gHg^{-1}$ be arbitrary. We need to show that $ghg^{-1} \in H$, i.e that there exists an $h' \in H$ so that $ghg^{-1} = h'$. But since $h \in H$ and $g^{-1} \in G$, the assumption guarantees us that $h \in g^{-1}Hg$. Therefore, we get that $g^{-1}hg = h'$ for some $h' \in H$ which implies that $h = gh'g^{-1}$ as desired. ■

The basic trick of the above proof is very common. Essentially, we used the fact that the G is the same set as the collection of all inverses of G.

WARNING! 8.46. This **does not** say that for a particular $g \in G$ that $gHg^{-1} \subseteq H$ implies that $g^{-1}Hg \subseteq H$. Indeed, consider the group of matrices $\mathbf{GL}_2(\mathbb{Q})$, the subgroup, H, to be

$$\left\{ \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} : n \in \mathbb{Z} \right\},$$

and the element g to be the matrix

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then, we have that

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2n \\ 0 & 1 \end{pmatrix}$$

which, since $2n$ is also an integer, is still an element of H. Hence we have that $gHg^{-1} \subseteq H$. However, we do not have equality (why?), and we do not have $g^{-1}Hg \subseteq H$. The latter is because

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$$

which is not an element of H.

If a subgroup H of G is normal we write $H \trianglelefteq G$. This is just a laziness technique that allows us to write less!

[III] This is very very rare for non-abelian groups!

WARNING! 8.47. We are **not** requiring that $ghg^{-1} = h$ for every $h \in H$. What we are really requiring is that for all $g \in G$ and $h \in H$ there exists some (possibly different) $h' \in H$ with $ghg^{-1} = h'$.

At this point it is very annoying (but pretty straight forward) to show that a subgroup of a group is normal by definition since we would have to check a case for every element of G and every element of H, a total of $|G| \cdot |H|$ cases .... zzzz. However, to show that a subgroup is not normal is much easier. We just need to find a g and a h so that Proposition 8.43 does not hold. When we begin talking about generators this will become much easier. As such we give only one example of a normal subgroup.

EXAMPLE 8.48.

(1) The subset $\{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\}$ is a subgroup of $\mathcal{D}_4$ since it is finite and closed under multiplication. Moreover, it is a normal subgroup of $\mathcal{D}_4$ since

$$\begin{aligned}
\mathbf{1}\{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\}\mathbf{1}^{-1} &= \{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\} \\
\rho\{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\}\rho^{-1} &= \{\mathbf{1}, \rho^2, \mathbf{r}\rho^2, \mathbf{r}\} \\
\rho^2\{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\}\rho^{-2} &= \{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\} \\
\rho^3\{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\}\rho^{-3} &= \{\mathbf{1}, \rho^2, \mathbf{r}\rho^2, \mathbf{r}\} \\
\mathbf{r}\{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\}\mathbf{r}^{-1} &= \{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\} \\
(\mathbf{r}\rho)\{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\}(\mathbf{r}\rho)^{-1} &= \{\mathbf{1}, \rho^2, \mathbf{r}\rho^2, \mathbf{r}\} \\
(\mathbf{r}\rho^2)\{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\}(\mathbf{r}\rho^2)^{-1} &= \{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\} \\
(\mathbf{r}\rho^3)\{\mathbf{1}, \rho^2, \mathbf{r}, \mathbf{r}\rho^2\}(\mathbf{r}\rho^3)^{-1} &= \{\mathbf{1}, \rho^2, \mathbf{r}\rho^2, \mathbf{r}\}.
\end{aligned}$$

(2) The subgroup $\langle\, \mathbf{r} \,\rangle = \{\mathbf{1}, \mathbf{r}\}$ is not normal in $\mathcal{D}_4$ since we have that

$$\rho\{\mathbf{1}, \mathbf{r}\}\rho^{-1} = \{\mathbf{1}, \rho\, \mathbf{r}\, \rho^{-1}\} = \{\mathbf{1}, \mathbf{r}\, \rho^2\} \neq \{\mathbf{1}, \rho\}.$$

Our first theoretical way of checking when a subgroup is normal – and the reason why the discussion of normal subgroups belongs in this chapter – is the following:

THEOREM 8.49. *The kernel of any homomorphism is a normal subgroup of its domain.*

PROOF. Let $\varphi\colon G \to H$ be a homomorphism between groups G and H. We have already seen that the kernel, $\operatorname{Ker}(\varphi)$ is a subgroup of G.

To finish the proof we must show that $\operatorname{Ker}(\varphi)$ is closed under conjugation by elements of G. Let $k \in \operatorname{Ker}(\varphi)$ and $g \in G$ be arbitrary. We want to show that

$$gkg^{-1} \in \operatorname{Ker}(\varphi)$$

as well. By definition, we are trying to show that $\varphi(gkg^{-1}) = e_H$. But

$$\begin{aligned}
\varphi(gkg^{-1}) &= \varphi(g)\varphi(k)\varphi(g)^{-1} && (\varphi \text{ is a homomorphism.}) \\
&= \varphi(g)e_H\varphi(g)^{-1} && (k \in \operatorname{Ker}(\varphi) \text{ by assumption.}) \\
&= \varphi(g)\varphi(g)^{-1} \\
&= e_H
\end{aligned}$$

as desired. ∎

Symbolically, this theorem says

$$\varphi\colon G \to H \text{ a gp hom } \Rightarrow \operatorname{Ker}(\varphi) \trianglelefteq G.$$

It will turn out that, in a sense, the converse of this statement is also true: every normal subgroup of G is the kernel of some homomorphism with domain G. We will prove this a few times, the first time being in Theorem 10.59 on page 528. Moreover, will show that if two homomorphisms have the same kernel, then their images are isomorphic. This will mean that in order to evaluate the possible homomorphisms from a group we just need to evaluate the set of normal subgroups of that group.

EXERCISE 8.33. Show that $\{e_G\}$ and G are normal subgroups of a group G. (The second statement says that any group is normal in itself).

SOLUTION

EXERCISE 8.34. Suppose we have a homomorphism of groups, $\varphi\colon G_1 \to G_2$.

(1) Suppose that $H_2 \trianglelefteq G_2$. Prove that

$$\varphi^{-1}(H_2) \trianglelefteq G_1.$$

Explain why this generalizes Theorem 8.49.

(2) Suppose that $H_1 \trianglelefteq G_1$. Show that it is not true in general (read find an example!) that

$$\varphi(H_1) \trianglelefteq G_2.$$

Explain why if it were true in general, that every subgroup of any group would be normal in that group.

(3) Suppose further that φ is surjective. Now prove that the hypothesis in the previous part is indeed true in this special case.

SOLUTION

EXERCISE 8.35. Show that the intersection of two normal subgroups is a normal subgroup.
Hint! Don't bother reproving the fact that the intersection of two subgroups is a subgroup. Just focus on proving the normal business.

SOLUTION

DEFINITION 8.50. A non-trivial group G is said to be *simple* if G does not contain a PNT, normal subgroup.

By Example/Theorem 6.26, for a prime number p, $\mathbb{Z}_p$ contains no PNT subgroups. Therefore, it cannot possibly contain a normal PNT subgroup, and it is therefore simple. The only other examples we have seen so far of simple groups are the groups $\mathbf{SL}_n(\mathbb{Z}_2)$ for $n > 2$ and $\mathcal{A}_n$ for $n > 4$. However, the proof of this fact is far from trivial. By the end of the course we will see two other families of simple groups. The definition of a simple group won't be used again until much later.

Exercise 8.36. Prove that a group G is simple iff every proper subgroup of G is malnormal.

Solution

Example /Theorem 8.51. If A is a finite, abelian simple group, then $A \cong \mathbb{Z}_p$ for some prime p.

Proof. Since every subgroup of an abelian group is normal, we only need to prove that if A is not isomorphic to $\mathbb{Z}_p$ that it contains a PNT subgroup. Let $a \in A$ be a non-trivial element (which exists since we excluded the trivial group as a simple group). Then $\langle a \rangle$ is a non-trivial subgroup of A. Since every subgroup of an abelian group is normal, we know that $\langle a \rangle \trianglelefteq A$. Since A is assumed simple, we must have that $\langle a \rangle = A$. In particular, $\langle a \rangle$ is cyclic (generated by the element a). But we've seen that a finite cyclic group contains a subgroup for every divisor of its order. Therefore A contains no proper subgroup iff $|A|$ is prime. ■

Example 8.52. For a commutative ring R, the set of invertible matrices of the form

$$\begin{pmatrix} x_{11} & x_{11} & x_{11} \\ 0 & x_{11} & x_{11} \\ 0 & 0 & x_{11} \end{pmatrix}$$

can easily checked to form a subgroup of $\mathbf{GL}_3(R)$ which we denote by $\mathrm{T}_3(R)$. For this same ring R, we also have the group subgroup of $\mathbf{GL}_3(R)$, $\mathrm{UT}_3(R)$ which we defined to be the set of matrices of the form

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}.$$

Since $\mathrm{UT}_3(R) \subseteq \mathrm{T}_3(R)$ as well, we have that $\mathrm{UT}_3(R) \preccurlyeq \mathrm{T}_3(R)$. We now show it's normal. Considering arbitrary elements

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{UT}_3(R) \quad \text{and} \quad \begin{pmatrix} x_{11} & x_{11} & x_{11} \\ 0 & x_{11} & x_{11} \\ 0 & 0 & x_{11} \end{pmatrix} \in \mathrm{T}_3(R).$$

Then a calculation that is just so much fun shows that

$$\begin{pmatrix} x_{11} & x_{11} & x_{11} \\ 0 & x_{11} & x_{11} \\ 0 & 0 & x_{11} \end{pmatrix} \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{11} & x_{11} & x_{11} \\ 0 & x_{11} & x_{11} \\ 0 & 0 & x_{11} \end{pmatrix}^{-1}$$
$$= \begin{pmatrix} 1 & -\frac{x_{12}}{x_{22}} + \frac{x_{11}a + x_{12}}{x_{22}} & \frac{x_{12}x_{23} - x_{13}x_{22}}{x_{22}x_{33}} - \frac{(x_{11}a + x_{12})x_{23}}{x_{22}x_{33}} + \frac{x_{11}b + x_{12}c + x_{13}}{x_{33}} \\ 0 & 1 & -\frac{x_{23}}{x_{33}} + \frac{x_{22}c + x_{23}}{x_{33}} \\ 0 & 0 & 1 \end{pmatrix}.$$

This matrix is still in $\mathrm{UT}_3(R)$ since there are 1's going down the diagonal (we don't give a crap about the nastiness above the diagonal!) Therefore we have proven that $\mathrm{UT}_3(R) \trianglelefteq \mathrm{T}_3(R)$.

Exercise 8.37. Show that the only subgroup of a group G which is both normal and malnormal is the trivial (sub)group.
Hint! This is as easy as you think it is!

SOLUTION

Exercise 8.38. We prove a few additional things about the core of a subgroup. This exercise can be thought of as saying that the core finds the biggest normal subgroup (of the bigger group) of a subgroup.

(1) For any subgroup $H \leqslant G$ show that $(H^\circ)^\circ = H^\circ$.

(2) Conclude that the core of any group is a normal subgroup.

(3) Show that if $N \subseteq H$ and $N \trianglelefteq G$, then $N \subseteq H^\circ$.

(4) Conclude that $H \leqslant G$ is malnormal iff H contains no non-trivial normal subgroup of G.

SOLUTION

Definition 8.53. The Klein 4–group is the set of four permutations

$$V \stackrel{\text{def}}{=} \{(1), (1\,2)(3\,4), (1\,3)(2\,4), (1\,4)(2\,3)\}. \tag{6}$$

Exercise 8.39. Prove that $V \trianglelefteq \mathcal{S}_4$.
Hint! I can think of several others ways of doing this besides performing the multitude of calculations the working definition would require. One uses the fact that the non-identity elements of V consist of all $\underline{2}\,\underline{2}$–cycles in $\mathcal{S}_4$. The other is to note that V is the kernel of some homomorphism $\mathcal{S}_4 \to \mathcal{S}_3$

SOLUTION

Definition 8.54. Let $H \leqslant G$. The *normalizer of H in G*, denoted $\mathrm{N}_G(H)$, is

$$\mathrm{N}_G(H) \stackrel{\text{def}}{=} \{g \in G : gHg^{-1} = H\}.$$

Exercise 8.40. We explore the normalizer further.

(1) Show that $\mathrm{N}_G(H) \leqslant G$.

SOLUTION

(2) Show that H is a normal subgroup of $\mathrm{N}_G(H)$.
Careful! Don't forget to show that H is actually a subset of its normalizer.

(3) Give an example of a group G and an $H \leqslant G$ where $\mathrm{N}_G(H) \not\trianglelefteq G$.

(4) Suppose that $H \leqslant K \leqslant G$ and that $H \trianglelefteq K$. Then show that $K \leqslant \mathrm{N}_G(H)$.[IV]

(5) Prove $H \trianglelefteq G \Leftrightarrow \mathrm{N}_G(H) = G$.

(6) Show that for a finite group H that we could instead define $\mathrm{N}_G(H)$ as

$$\{g \in G : \forall h \in H\, ghg^{-1} \in H\}$$

(i.e. this is the set of $g \in G$ with $gHg^{-1} \subseteq H$).

We say that $H \leqslant G$ is *self–normalizing* if $\mathrm{N}_G(H) = H$.

Exercise 8.41. Given a copy, $\widetilde{\mathcal{D}_4}$, of $\mathcal{D}_4$ inside of $\mathcal{S}_4$ obtained by labeling the vertices of a square with the numbers 1 through 4,[V] prove that $\widetilde{\mathcal{D}_4}$ is self–normalizing but is not malnormal.
Hint! To prove that $\widetilde{\mathcal{D}_4}$ is not malnormal, note that it contains the Klein four group, V, from (6). At this point (later, this will be trivial to prove), to prove that $\widetilde{\mathcal{D}_4}$ is self–normalizing requires a bit of ingenuity. I would use the fact that for σ in the normalizer, we have that

$$\sigma(1\,2\,3\,4)\sigma^{-1} = (1\,2\,3\,4) \quad \text{or} \quad \sigma(1\,2\,3\,4)\sigma^{-1} = (1\,4\,3\,2)$$

(since these are the only $\underline{4}$–cycles in $\widetilde{\mathcal{D}_4}$. This means that either σ or $(2\,4)\sigma$ is in the centralizer of $(1\,2\,3\,4)$. Then use the Exercise 7.6.

SOLUTION

Exercise 8.42. Show that being normal isn't transitive (so $\trianglelefteq$ is not a normal relation but $\leqslant$ is). That is find subgroups $H \leqslant K \leqslant G$ with $H \trianglelefteq K$ and $K \trianglelefteq G$ but with $H \not\trianglelefteq G$.

SOLUTION

Exercise 8.43. Show that if $N \trianglelefteq G$ and $H \leqslant G$, then $H \cap N \trianglelefteq H$. Also prove that this is false if we replace normal by malnormal. Show that if H is malnormal in G, then so is any subgroup of H. (This implies that being malnormal is transitive in a very strong way!)

SOLUTION

The following exercise gives us another handy dandy way of finding normal subgroups. If we take $H = \mathcal{C}(G)$ it shows that the center of any group is normal.

Exercise 8.44. If $H \leqslant \mathcal{C}(G)$, then $H \trianglelefteq G$.

SOLUTION

EXAMPLE /THEOREM 8.55. The normalizer in $\mathbf{GL}_n\left(\mathbb{Z}_p\right)$ of $\mathrm{UT}_n\left(\mathbb{Z}_p\right)$ is $\mathrm{T}_n\left(\mathbb{Z}_p\right)$ and in $\mathbf{SL}_n\left(\mathbb{Z}_p\right)$ it's

$$\mathrm{ST}_n\left(\mathbb{Z}_p\right) \stackrel{\text{def}}{=} \mathrm{T}_n\left(\mathbb{Z}_p\right) \cap \mathbf{SL}_n\left(\mathbb{Z}_p\right)$$

where $\mathrm{T}_n\left(\mathbb{Z}_p\right)$ is defined to be the group of invertible upper triangular matrices.

PROOF. We need to prove two inclusions to prove that

$$\mathrm{N}_{\mathbf{GL}_n(\mathbb{Z}_p)}\mathrm{UT}_n\left(\mathbb{Z}_p\right) = \mathrm{T}_n\left(\mathbb{Z}_p\right).$$

The statement about the normalizer of the subgroup in $\mathbf{SL}_n\left(\mathbb{Z}_p\right)$ follows in general since

$$\mathrm{N}_H(K) = \mathrm{N}_G(K) \cap H$$

whenever $H \leqslant G$.

$\subseteq$ Of course we could do a direct and hideous calculation of this, but this would be rather nasty. Let

$$\mathbf{A} \in \mathrm{N}_{\mathbf{GL}_n(\mathbb{Z}_p)}\left(\mathrm{UT}_n\left(\mathbb{Z}_p\right)\right)$$

be arbitrary. And let $\mathbf{L}$ denote the matrix which consists only of 1's on the diagonal and 1 above the diagonal (and zeros elsewhere) Then, by definition of the normalizer,

$$\mathbf{A}^{-1}\mathbf{L}\mathbf{A} = \mathbf{L}' \Rightarrow \mathbf{L}\mathbf{A} = \mathbf{A}\mathbf{L}'$$

where $\mathbf{L}'$ is some element of $\mathrm{UT}_n\left(\mathbb{Z}_p\right)$. Then since the only vector $\mathbf{L}$ fixes are multiples of $\mathbf{e}_1$, and every element of $\mathrm{UT}_n\left(\mathbb{Z}_p\right)$, including $\mathbf{L}'$, fixes $\mathbf{e}_1$, we know that

$$\mathbf{L} \cdot (\mathbf{A} \cdot \mathbf{e}_1) = \mathbf{A} \cdot (\mathbf{L}' \cdot \mathbf{e}_1) \Rightarrow \mathbf{L} \cdot (\mathbf{A} \cdot \mathbf{e}_1) = \mathbf{A}(\mathbf{e}_1).$$

Therefore the matrix $\mathbf{L}$ fixes $\mathbf{A} \cdot \mathbf{e}_1$, and as stated, this means that $\mathbf{A} \cdot \mathbf{e}_1 = d_{11}\mathbf{e}_1$ where d_{11} is some (non-zero) scalar (i.e. element of $\mathbb{Z}_p$). For ease of notation, we denote the matrix $\mathbf{A}/d_{11}$ as $\mathbf{A}_1$. As a multiple of $\mathbf{A}$, we still have that $\mathbf{A}_1$ normalizes $\mathrm{UT}_n\left(\mathbb{Z}_p\right)$. Moreover $\mathbf{A}_1$, unlike $\mathbf{A}$, fixes $\mathbf{e}_1$.

Next note that

$$(\mathbf{L}' - \mathbf{I}_n) \cdot \mathbf{e}_2 = c_1\mathbf{e}_1$$

where c_1 is some scalar, and

$$(\mathbf{L} - \mathbf{I}_n)(\mathbf{v}) = \mathbf{e}_1 \Leftrightarrow \mathbf{v} = \mathbf{e}_2 + \alpha_2\,\mathbf{e}_1.$$

for some $\alpha_2 \in \mathbb{Z}_p$. Therefore

$$\begin{aligned}
c_1\mathbf{e}_1 &= \mathbf{A}_1(c_1\,\mathbf{e}_1) \\
&= \mathbf{A}_1 \cdot ((\mathbf{L}' - \mathbf{I}_n) \cdot \mathbf{e}_2) \\
&= (\mathbf{L} - \mathbf{I}_n)(\mathbf{A}_1 \cdot \mathbf{e}_2) \\
&\Rightarrow \mathbf{e}_1 = \left(\mathbf{L} - \mathbf{I}_n\right) \cdot \left(\frac{\mathbf{A}_1}{c_2} \cdot \mathbf{e}_2\right) \\
&\Rightarrow \left(\frac{\mathbf{A}_1}{c_2} \cdot \mathbf{e}_2\right) = \mathbf{e}_2 + \alpha_2\,\mathbf{e}_1, \text{ for some } \alpha_2 \in \mathbb{Z}_p \\
&\Rightarrow \mathbf{A}(\mathbf{e}_2) = d_{12}\,\mathbf{e}_1 + d_{22}\mathbf{e}_2,\ d_{12}, d_{22} \in \mathbb{Z}_p.
\end{aligned}$$

for some $\alpha_2 \in \mathbb{Z}_p$. {Finish }

$\supseteq$

■

8.5 Isomorphism Invariants

In a sense, two isomorphic groups cannot be distinguished by any group theoretical property. Therefore, it is in this sense, that two isomorphic groups in a group theory class are considered to be "the same." It's typically an easier problem to show that two groups are not isomorphic than it is to show that they are not. What is the converse of what it means for two groups to be isomorphic? Well, we are required that every bijection between them does not preserve operations (i.e. is not a group homomorphism). This is clearly not a practical method of showing this. It turns out to be considerably easier to find a list of certain properties that any two isomorphic groups have in common. Such properties are known as *isomorphism invariants*. Roughly speaking, nearly everything that can be stated about a group that "sounds" like it doesn't depend on how you name elements will be an isomorphism invariant. For instance, since an isomorphism is a bijection, any two isomorphic groups must have the same number of elements. This shows, without a doubt, that the groups $\mathcal{S}_4$ and $\mathcal{D}_4$ are not isomorphic since the former has $4! = 24$ elements while the latter has $4 \cdot 2 = 8$ elements in it. Also if a and b commute in G and φ is a homomorphism, then

$$\varphi(a)\varphi(b) = \varphi(ab) = \varphi(ba) = \varphi(b)\varphi(a)$$

implies that their images also commute. Since isomorphisms are surjective this proves that isomorphic groups are either both abelian or they are both not abelian. Other isomorphism invariants include the number of elements of a given (finite) order (Theorem 8.22), the size of the center (easy adaptation of Exercise 8.21), the number of subgroups of a certain order, the possible isomorphism types of subgroups, the possible sizes of conjugacy classes of elements, and many many more.

More rigorously, a property of groups is an equivalence relation defined on the class of all groups. If such a property contains the equivalence relation which relates two isomorphic groups, then this property is an isomorphism invariant. In other words if $\sim$ is such a property, then

$$G \cong H \Rightarrow G \sim H.$$

(Of course this is often akward, but it's an easy way to state the definition formally!) For instance, in this bizarre way, the order of a group is a property which relates two groups iff they have the same number of elements. **Abelian** is also a property which relates two groups iff they are abelian.

Example 8.56.

(1) Even though they both have order 6, the groups $\mathcal{S}_3$ and $\mathbb{Z}_6$ are not isomorphic. This is because $\mathcal{S}_3$ is non-abelian while $\mathbb{Z}_6$ is abelian.

(2) Even though they both have the same order and are both abelian, the two groups $\mathbb{Z}_4$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$ are non-isomorphic. This is because $\mathbb{Z}_4$ contains an element of order 4 while the second group does not.

(3) Even though they have the same order and both are non-abelian, the groups $\mathcal{D}_4$ and $\mathcal{Q}_8$ are not isomorphic. This is because $\mathcal{Q}_8$ contains only one element of order 2, namely the quaternion -1 while $\mathcal{D}_4$ has more than 1 (actually it has 5). For instance any reflection has order 2 and is not a 180° rotation, which also has order 2.

Exercise 8.45. Prove that the groups $\mathcal{A}_4$ and $B\mathcal{D}_3$ are not isomorphic.

Solution

Often one cannot get away with simply using these simple binary or numerical invariants. The following exercise illustrates this.

Exercise 8.46. Show that $\mathcal{S}_3$ is not isomorphic to any subgroup of $\mathcal{A}_4$.
Hint! Show that any two elements of order 2 in $\mathcal{A}_4$ commute and no two elements of order 2 in $\mathcal{S}_3$ commute. Be sure to explain why this difference guarantees that the statement of the exercise is true.

Solution

DEFINITION 8.57. For any group G, $\mathbf{o}_n(G)$ denotes the number of elements of order n in G. The *order table* of a finite group is simply a table which contains the number of elements of a given order.

- We will always write these tables in increasing order.
- We will see as a consequence of Lagrange's theorem, Corollary 10.69.1 back on page 532, that the order of every element of a finite group divides the order of the group. Thus this table only needs to consider numbers which divide the order of the group. We will not officially "use" this fact until we end up proving it, but we may appear "lazy" when discussing the possible orders elements of a given group can have.

We will always present order tables in a manor similar to what we present in the below examples.

(1) The group $\mathbb{Z}_6$ has order table

n	1	2	3	6
$\mathbf{o}_n$	1	1	2	2

.

(2) The group $\mathcal{S}_3$ has order table

n	1	2	3	6
$\mathbf{o}_n$	1	3	2	0

since $\mathcal{S}_3$ (like every group) contains one element of order 1, namely the identity, 3 elements of order 2, namely $(1\,2)$, $(2\,3)$, and $(1\,3)$, and 2 elements of order 3, namely $(1\,2\,3)$ and $(1\,3\,2)$.

(3) The group $\mathcal{D}_4$ has order table

n	1	2	4
$\mathbf{o}_n$	1	5	2

.

(4) The group $\mathcal{Q}_8$ has order table

n	1	2	4
$\mathbf{o}_n$	1	1	6

.

(5) In Exercise 5.41 you showed that the order table of $\mathrm{UT}_3(\mathbb{Z}_3)$ has the following simple form

n	1	3
$\mathbf{o}_n$	1	26

.

(6) We use a bit of trickery to figure out the order table of say a group like $\mathcal{S}_3 \times \mathcal{S}_3$. Clearly we only need to figure out the number of elements of order 1,2,3 or 6. We use Proposition 5.94 back on page 297 to help us.

However we know by now that these numbers are not the number of elements in the group since, of course, the order is the least power which is the identity. However we can use an inclusion/exclusion argument to get our answers. For instance, to figure out the number of elements of order 6, in a group G, $\mathbf{o}(G)$, we can simply compute

$$\mathbf{o}_6(G) = \mathbf{O}_6(G) - \mathbf{O}_3(G) - \mathbf{O}_2(G) + \mathbf{O}_1(G).$$

Now we know that

$$\mathbf{O}_6(\mathcal{S}_3) = 6 \qquad \mathbf{O}_3(\mathcal{S}_3) = 3 \qquad \mathbf{O}_2(\mathcal{S}_3) = 4 \qquad \mathbf{O}_1(\mathcal{S}_3) = 1.$$

Therefore we have

$$\mathbf{O}_6(\mathcal{S}_3 \times \mathcal{S}_3) = 6^2 = 36 \qquad \mathbf{O}_3(\mathcal{S}_3 \times \mathcal{S}_3) = 3^2 = 9$$
$$\mathbf{O}_4(\mathcal{S}_3 \times \mathcal{S}_3) = 2^2 = 4 \qquad \mathbf{O}_1(\mathcal{S}_3 \times \mathcal{S}_3) = 1^2 = 1.$$

So we have

$$\mathbf{o}_6(\mathcal{S}_3 \times \mathcal{S}_3) = 36 - 16 - 9 + 1 = 12.$$

Meanwhile,

$$\mathbf{o}_3(\mathcal{S}_3 \times \mathcal{S}_3) = \mathbf{O}_3(\mathcal{S}_3 \times \mathcal{S}_3) - \mathbf{O}_1(\mathcal{S}_3 \times \mathcal{S}_3) = 9 - 1 = 8,$$

and

$$\mathbf{o}_3(\mathcal{S}_3 \times \mathcal{S}_3) = \mathbf{O}_2(\mathcal{S}_3 \times \mathcal{S}_3) - \mathbf{O}_1(\mathcal{S}_3 \times \mathcal{S}_3) = 16 - 1 = 15.$$

In summary we have the order table for $\mathcal{S}_3 \times \mathcal{S}_3$ is

$\mathcal{S}_3 \times \mathcal{S}_3$	n	1	2	3	6
	$\mathbf{o}_n$	1	15	8	12

.

THEOREM 8.58. *Two isomorphic groups have the same order table.*

PROOF. An isomorphism φ takes elements of a given order to elements of that same order. ∎

EXAMPLE 8.59. We are about to, in Theorem 8.61 coming up on page 419, prove that any two cyclic groups of the same order are isomorphic. Since 31 is a prime number, the group $\mathcal{U}(\mathbb{Z}_{31})$ is cyclic. Therefore, as cyclic groups of order 30, the groups $\mathcal{U}(\mathbb{Z}_{31})$ and $\mathbb{Z}_{30}$ are isomorphic to each other. As such, have the same order table. They are both (Here we used the fact that for a divisor d of 30 there are $\vartheta(d)$ elements of order d in a cyclic group of order 30.)

n	1	2	3	5	6	10	15	30
$\mathbf{o}_n$	1	1	2	4	2	4	8	8

Of course this wouldn't be very much fun to do directly for the group $\mathcal{U}(\mathbb{Z}_{31})$!

EXERCISE 8.47. Explicitly find the 2 elements of order 3 in $\mathcal{U}(\mathbb{Z}_{31})$.

SOLUTION

The following exercise shows that the converse is false: there are non-isomorphic groups with the same order table. (The first example of this occurs for two groups of order 27.)

EXERCISE 8.48. Show that $\mathbb{Z}_3^3$ and $UT_3(\mathbb{Z}_3)$ (where the groups, $UT_3(\mathbb{Z}_n)$, were defined in Exercise 6.25 back on page 335) have the same order table but are not isomorphic groups.

SOLUTION

EXERCISE 8.49. Calculate the order tables of $\mathbb{Z}_3 \times \mathcal{S}_3$ and $\mathcal{D}_9$ (both groups are non-abelian groups of order 18). Show that these groups are not isomorphic (of course you can simply show that $\mathcal{D}_9$ contains an element of order 9 while $\mathcal{S}_3 \times \mathcal{S}_3$ does not).

SOLUTION

EXERCISE 8.50. Suppose that $H \preccurlyeq G$. Prove that for every integer m, that $\mathbf{o}_m(H) \leq \mathbf{o}_m(G)$.
Hint! Like many questions I like to ask this is simply a question of whether or not you understand the notation.

SOLUTION

EXERCISE 8.51. Prove that the groups $\mathcal{D}_3 \times \mathbb{Z}_{11}$ and $\mathcal{D}_{11} \times \mathbb{Z}_3$ are two non-isomorphic groups.

SOLUTION

8.5.1 Groups of order 24.

The groups $\mathcal{S}_4$ and $\mathcal{D}_{12}$ are not isomorphic (non–abelian groups of order 24). While we could check that $\mathcal{S}_4$ does not contain an element of order 12 but $\mathcal{D}_{12}$ does, we could also note that $\mathcal{D}_{12}$ is not centerless:

$$180^\circ \text{ rot} = \rho_{12}^6 \in \mathcal{C}(\mathcal{D}_{12}).$$

On the other hand, in the previous chapter, in Exercise 7.4, you showed that $\mathcal{S}_n$ (and hence $\mathcal{S}_4$) is centerless for $n > 2$. Meanwhile, the group $B\left(\blacktriangleright^3\right)$ contains only 1 element of order 2 while both of the other groups, $\mathcal{S}_4$ and $\mathcal{D}_{12}$ have more than 1 element of order 2, thus we have found a third isomorphism type of a non–abelian group of order 24. Let's keep going? Next $B\mathcal{D}_6$ is another non-abelian group of order 24 which contains only 1 element of order 2. But it can't be isomorphic to $B\left(\blacktriangleright^3\right)$ since $B\left(\blacktriangleright^3\right)$ does not contain an element of order 12 while $B\mathcal{D}_6$ does, namely χ_{12}. Next, $\mathcal{D}_6 \times \mathbb{Z}_2$ cannot be isomorphic to any of the groups mentioned, so far since it's center contains four elements (and is isomorphic to $\mathbb{Z}_2^2$) while every other group mentioned so far contains at most 2 elements in its center. Chugging along, $B\mathcal{D}_3 \times \mathbb{Z}_4$ and $\mathcal{D}_3 \times \mathbb{Z}_2^2$ are not isomorphic to any group listed or each other. The reason being that $B\mathcal{D}_3 \times \mathbb{Z}_4$ has a center isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_4$ and $\mathcal{D}_3 \times \mathbb{Z}_2^2$ has a center isomorphic to $\mathbb{Z}_2^3$.

EXERCISE 8.52. Technically in these problems I am simply asking you to rule out all but one of the groups from the table.

(1) We will see later that the group $\mathbf{SL}_2(\mathbb{Z}_3)$ is isomorphic to one of the groups in Table 8.2. Which one will it have to be?

(2) We will see later that the group $\mathcal{S}_3 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ is isomorphic to one of the groups in Table 8.2. Which one will it have to be?

SOLUTION

G	$\mathcal{S}_4$	$B\left(\blacktriangleright^3\right)$	$B\mathcal{D}_6$	$\mathcal{D}_{12}$	$\mathcal{D}_6^{\times 2}$	$\mathcal{Q}_8^{\times 3}$	$\mathcal{D}_4^{\times 3}$	$\mathcal{A}_4^{\times 2}$	$B\mathcal{D}_3^{\times 2}$	$\mathcal{S}_3^{\times 4}$
LPO(G)	4	6	12	12	6	12	12	6	6	12
$\mathcal{C}(G)$	e	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}_2^2$	$\mathbb{Z}_6$	$\mathbb{Z}_6$	$\mathbb{Z}_2$	$\mathbb{Z}_2^2$	$\mathbb{Z}_4$
$\mathbf{o}_2(G)$	9	1	1	13	15	1	5	7	3	7

Table 8.1: Some Non-Abelian Groups of Order 24. The entry in the LPO row is the largest order an element of the corresponding group. To make the table fit better I use the notation that for a group G, $G^{\times n}$ means $G \times \mathbb{Z}_n$. I only use this notation to save space!

Cheating and using Appendix I on page 865 we see that we are still missing to (isomorphism types) of groups of order 24. At this point the two other groups seem a bit difficult to describe. We will be able to do this later in Section 12.5.1 coming up on page 673. If we call the two groups missing G_{24}^1 and G_{24}^2, then we'll see that the table will be completed as follows:

G	$\mathcal{S}_4$	$B\left(\blacktriangleright^3\right)$	$B\mathcal{D}_6$	$\mathcal{D}_{12}$	$\mathcal{D}_6^{\times 2}$	$\mathcal{Q}_8^{\times 3}$	$\mathcal{D}_4^{\times 3}$	$\mathcal{A}_4^{\times 2}$	$B\mathcal{D}_3^{\times 2}$	$\mathcal{S}_3^{\times 4}$	G_{24}^1	G_{24}^2
LPO(G)	4	6	12	12	6	12	12	6	6	12	12	6
$\mathcal{C}(G)$	e	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}_2^2$	$\mathbb{Z}_6$	$\mathbb{Z}_6$	$\mathbb{Z}_2$	$\mathbb{Z}_2^2$	$\mathbb{Z}_4$	$\mathbb{Z}_4$	$\mathbb{Z}_2$
$\mathbf{o}_2(G)$	9	1	1	13	15	1	5	7	3	7	1	9

Table 8.2: All Non–Abelian Groups of Order 24 The two groups G_{24}^1 and G_{24}^2 introduced on 674.

EXERCISE 8.53. Prove that $B\mathcal{D}_6$ does not contains an element of order 8.
Hint! If it did you could use the homomorphism

$$B\mathcal{D}_6 \to B\mathcal{D}_6$$

to prove that $B\mathcal{D}_6$ has an element of order 4 or 8. Then use Example/Theorem 5.108 back on page 304.

SOLUTION

WARNING! 8.60. We have **not** proven that these are all the possible groups of order 24! That is quite difficult to do! All we have shown is that all the groups listed are not isomorphic! This section was made to give you a taste of one of the first numbers where there are a wide assortment of interesting groups.

8.6 Kernels Uniquely Determine Images

Let G be a group and $g \in G$. To motivate this section, recall from Corollary 5.62 back on page 282 that if $g \in G$ has order n (we are allowing $n = \infty$), then

$$\begin{aligned} \mathbf{E}_g^n &: \mathbb{Z}_n \to \langle\, g \,\rangle \\ &: [\, a \,]_n \mapsto g^a \end{aligned}$$

was a well defined bijection. Since we have a bijective function we simply need Now to see that this is an isomorphism is now trivial; we just have to show that the map $\mathbf{E}_g^n$ is operation preserving. Indeed, this proof should "feel" just like the proof which showed that the maps $\mathbf{E}_g : \mathbb{Z} \to G$ were homomorphisms.

$$\begin{aligned} \mathbf{E}_g^n([\, a \,]_n + [\, b \,]_n) &= \mathbf{E}_g^n([\, a + b \,]_n) \\ &= g^{a+b} \\ &= g^a\, g^b \\ &= \mathbf{E}_g^n([\, a \,]_n) \cdot \mathbf{E}_g^n([\, b \,]_n). \end{aligned}$$

Thus we have proven the following nice result: (For consistency, in the theorem below, we set $\mathbb{Z}_\infty = \mathbb{Z}_0 = \mathbb{Z}$ and $\mathbf{E}_g^0 = \mathbf{E}_g$.)

Theorem 8.61.

(1) *A cyclic group of order n is isomorphic to $\mathbb{Z}_n$.*

Or in more detail:

(2) *Let g be any element of any group G. Suppose that $|\, g \,| = n$. Then $\mathbb{Z}_n \cong \langle\, g \,\rangle$ via the isomorphism*

$$\mathbf{E}_g^n : \mathbb{Z}_n \to \langle\, g \,\rangle.$$

■

For example, since the roots of unity groups, $\{\{n\}\}$, are all cyclic (generated by $\chi_n = e^{\frac{2\pi i}{n}}$) we have

Corollary 8.62. *For any positive integer n the groups $\{\{n\}\}$ and $\mathbb{Z}_n$ are isomorphic.* ■

In fact we started the chapter with this special case where we showed the isomorphism $\mathbb{Z}_4 \cong \{1, -1, i, -i\}$ by directly comparing tables. Another special case of Theorem 8.61 is

Corollary 8.63. *The group*

$$\mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_m}$$

is isomorphic to $\mathbb{Z}_{n_1 \cdots n_m}$ iff the n_i are pair-wise relatively prime. ■

As a special case (the special case of this corollary when $m = 2$) we have

Corollary[2] 8.64. *The groups*

$$\mathbb{Z}_n \times \mathbb{Z}_m \cong \mathbb{Z}_{nm} \Leftrightarrow \gcd(n, m) = 1.$$

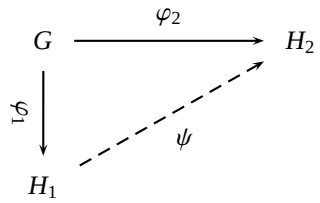
■

As was the case here, the bulk of the proof was actually showing that the map $\mathbf{E}_g^n$ was well-defined. This was only obvious in the case where $|\, g \,|$ is infinite and so the domain is $\mathbb{Z}$.

Here is a very general theorem that will be useful later. For now we use a very very special case of it.

THEOREM 8.65. *If two group homomorphisms have the same domain and the same kernel, then they have isomorphic images.*

First a lemma:



LEMMA 8.66. *Suppose that φ_1 and φ_2 are group homomorphisms from a group G to groups H and K (resp). Then there exists a group homomorphism ψ from $\varphi_1(G) \leqslant H$ to K which satisfies $\psi \circ \varphi_1 = \varphi_2$ iff*

$$\operatorname{Ker}(\varphi_1) \subseteq \operatorname{Ker}(\varphi_2).$$ [I]

Moreover when ψ exists we also have that

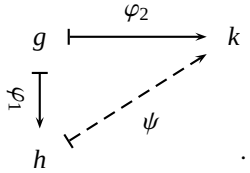
(1) $\operatorname{Im}(\psi) = \operatorname{Im}(\varphi_1)$*;*

(2) $\operatorname{Ker}(\psi) = \varphi_1(\operatorname{Ker}(\varphi_2))$

PROOF. By replacing H with $\varphi_1(G)$ and K with $\varphi_2(G)$ we can assume that both φ_1 and φ_2 are surjective.

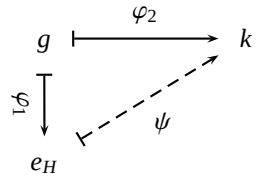
We define the relation ψ as the relation $\varphi_2 \circ \varphi_1^{-1}$.[II] Using some basic properties of relations we can see that there exists a function ψ' so that $\psi' \circ \varphi_1 = \varphi_2$ iff ψ is a function and in which case $\psi = \psi'$.

As usual, we use function notation for ψ even though we only know that its a relation and we do not yet know that it's a function. It's useful to think of ψ as saying that $\psi(h) = k$ iff there exists a $g \in G$ with



(Meaning $\varphi_1(g) = h$ and $\varphi_2(g) = k$.) [III]

$\Leftarrow$ (We prove the contrapositive) Suppose that $\operatorname{Ker}(\varphi_1)$ is not a subset of $\operatorname{Ker}(\varphi_2)$. There there exists a $g \in G$ which is in the first set, $\operatorname{Ker}(\varphi_1)$ but which is not in the second. That is $\varphi_1(g) = e_H$ but $k := \varphi_1(g) \neq e_K$. But then

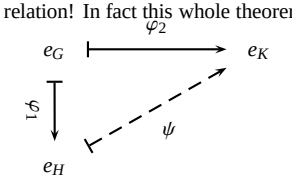


we have that $\psi(e_H) = k \neq e_K$ which is impossible if ψ is to be a group homomorphism.[IV]

[I] Written out, this means that $\varphi_2(g) = e_H$ whenever $\varphi_1(g) = e_K$.

[II] We are **not** assuming that φ_1 is a bijection, therefore all we know about φ_1^{-1} is that it is a relation! In fact this whole theorem is obvious and pointless when φ_1 is a bijection, i.e. when φ_1^{-1} is a function.

[III] For instance since we have



we know that $\psi(e_H) = e_K$.

[IV] Actually, this really doesn't have anything to do with the fact that we are dealing with groups since we've already proved that $\psi(e_H) = e_K$ in the previous footnote.

$\Rightarrow$ Conversely, we now suppose that

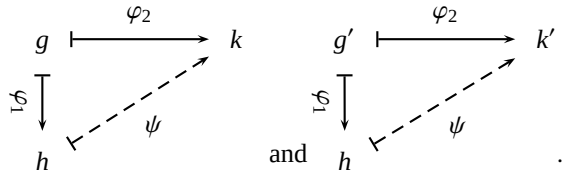
$$\mathrm{Ker}\,(\varphi_1) \subseteq \mathrm{Ker}\,(\varphi_2)\,.$$

and we now prove that ψ is a function.

(a) (ψ **is well defined**) Suppose that we have

$$\psi(h) = k \quad \text{and} \quad \psi(h) = k'.$$

We need to show that $k = k'$. In terms of diagrams we know there exists $g, g' \in G$ with



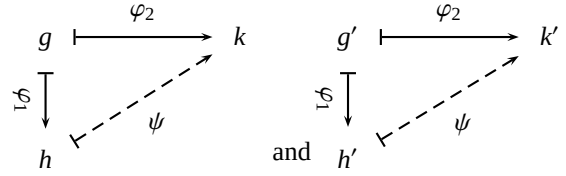
But we have

$$\begin{aligned} \varphi_1(g) = \varphi_1(g') &\Leftrightarrow g^{-1}g' \in \mathrm{Ker}\,(\varphi_1) \quad \text{(See footnote, [V], below)} \\ &\Rightarrow g^{-1}g' \in \mathrm{Ker}\,(\varphi_2) \quad \text{(Assumption)} \\ &\Leftrightarrow \varphi_2(g) = \varphi_2(g'). \quad \text{([V] again)} \end{aligned}$$

But $\varphi_1(g) = \varphi_1(g')$ is indeed true since both equal h. Therefore by our logic we also have that $\varphi_2(g) = \varphi_2(g')$. But the lhs is k and the rhs is k'.

Having shown that ψ is well defined, the rest is pretty easy.

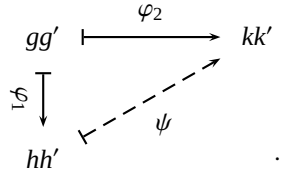
(b) (ψ **is a operation preserving**) Let h and $h' \in H$. Write $k := \psi(h)$ and $k' := \psi(h')$. We need to prove that $\psi(hh') = kk'$. We have diagrams



for some $g, g' \in G$. All this means is that $\varphi_1(g) = h$ and $\varphi_1(g') = h'$ while at the same time $\varphi_2(g) = k$ and $\varphi_2(g) = k'$.
But then since we are assuming that both φ_1 and φ_2 are group homomorphism,

$$\varphi_1(gg') = \varphi_1(g)\varphi_1(g') = hh' \text{ and } \varphi_2(gg') = \varphi_2(g)\varphi_2(g') = kk'.$$

In other words we have the diagram

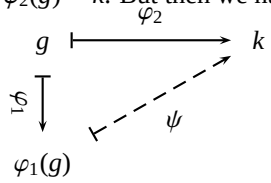


Therefore $\psi(hh') = kk'$ as desired.

[V] For any group homomorphism α,

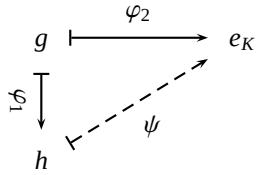
$$\alpha(g) = \alpha(g') \Leftrightarrow \alpha(g)^{-1}\alpha(g') = e_G \Leftrightarrow \alpha(g)^{-1}\alpha(g') \Leftrightarrow \alpha(g^{-1}g') = e_G \Leftrightarrow g^{-1}g' \in \mathrm{Ker}\,(\alpha)$$

(c) (ψ **is surjective**) [This part works in general, and again, has nothing to do with group theory or algebra of any kind.] Let $k \in K$ be arbitrary. Then, since we are assuming that $\varphi_2 : G \to K$ is surjective, there exists a $g \in G$ so that $\varphi_2(g) = k$. But then we have



and so $\varphi_1(g)$ gets mapped to k under ψ. Since k was arbitrary, ψ is surjective.

(d) ($\text{Ker}\psi = \varphi_1(\text{Ker}\varphi_2)$) We have



iff $\varphi_1(g) = h$ and $\varphi_2(g) = e_K$, i.e. $g \in \text{Ker}(\varphi_2)$. Therefore $h \in \text{Ker}(\psi)$ iff there exists a $g \in \text{Ker}(\varphi_2)$ with $\varphi_1(g) = h$. So

$$\text{Ker}(\psi) = \{\varphi(g) : g \in \text{Ker}(\varphi_2) = \varphi_1(\text{Ker}(\varphi_2))$$

as desired. ∎

Definition 8.67. If the homomorphism ψ exists, we call it the induced homomorphism.

Remark 8.68. If such a homomorphism, ψ, exists it must be defined by the relation $\varphi_2 \circ \varphi_1^{-1}$ considered in the proof. This is because we would have

$$\begin{aligned} \psi \circ \varphi_1 &= \varphi_2 \\ \Rightarrow \psi &= \psi \circ (\varphi_1 \circ \varphi_1^{-1}) \\ &= (\psi \circ \varphi_1) \circ \varphi_1^{-1} \\ &= \varphi_2 \circ \varphi_1^{-1}. \end{aligned}$$

Since in general φ_1^{-1} is not a function, these are equalities of relations, not functions! Despite the fact that φ_1^{-1} is often not a function, we have derived a condition for $\varphi_2 \circ \varphi_1^{-1}$ to be a function. Here we are using what we proved in Example 2.37 way back page 25.

Proof. (Proof of Theorem 8.65) Let $\varphi_1 : G \to H$ and $\varphi_2 : G \to K$ be two homomorphisms with the same kernel, $N \trianglelefteq G$. By replacing images by codomains, we can assume that both φ_1 and φ_2 are surjective (this only amounts to a notational change!). Note that this assumption does not change the kernel. Since the assumption of the theorem is that $\text{Ker}(\varphi_1) = \text{Ker}(\varphi_2)$, we have, in particular that $\text{Ker}(\varphi_1) \subseteq \text{Ker}(\varphi_2)$. Hence the lemma applies to give us an epimorphism of groups $\psi : H \to K$ which satisfies $\psi \circ \varphi_1 = \varphi_2$.

We need to prove that ψ is also injective. To accomplish this, we apply our favorite lemma and simply prove that the kernel is trivial. Indeed, assume that $h \in \text{Ker}(\psi)$. Then $\psi(h) = e_G$. Since we are assuming that φ_1 is surjective, we have an $g \in G$ so that $\varphi_1(g) = h$. But this means that

$$\varphi_2(g) = \psi(\varphi_1(g)) = \psi(h) = e_K.$$

In other words $g \in \mathrm{Ker}(\varphi_2)$. But we are assuming that the two sets $\mathrm{Ker}(\varphi_1) = \mathrm{Ker}(\varphi_2)$. Hence we must also have that $g \in \mathrm{Ker}(\varphi_1)$, i.e. $\varphi_1(g) = e_H$ as well. But, we defined h as $\varphi_1(g)$. Hence we have proven that $h = e_H$. Since h was an arbitrary element of $\mathrm{Ker}(\psi)$ we have proven that $\mathrm{Ker}(\psi) = e_H$. Hence the kernel is trivial and ψ is injective. ■

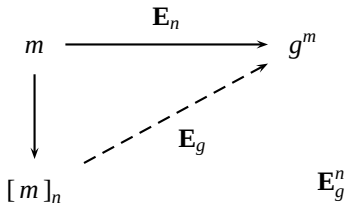
REMARK 8.69. Proofs along these lines are called "diagram chases." This proof is (probably) the simplest example of a diagram chase! Perhaps it would be useful for you to draw the draw consisting of G, H, and K like I did above and continuously follow how we pushed the elements around. While this type of proof is initially quite mystifying, after one gets used to them they become notoriously left brained! There are many puns about these types of proofs out there including the pun that such proofs should never be done in public! Wiki, agreeing with me, states that

> Although it may boggle the mind at first, once one has some practice at it, it is actually fairly routine.

We repeat Theorem 8.61 to help highlight how to use the above theorem. Notice that in the proof below, unlike the previous proof of this, we only use the theorem and no other facts. In particular, we don't need to know anything about the maps $\mathbf{E}_g$ and $\mathbf{E}_g^n$ presented earlier in this chapter or in Chapter 5. If we set $\mathbb{Z}_\infty = \mathbb{Z}$ the following corollary holds even for $n = \infty$.

COROLLARY 8.70. *(Repeat of Theorem 8.61) Let g be any element of any group G. Suppose that $|g| = n$. Then $\langle g \rangle \cong \mathbb{Z}_n$.*

PROOF. In the following diagram we have that both maps



$m \mapsto [\,m\,]_n$ and $m \mapsto g^m$ have kernel $n\mathbb{Z}$ and have images $\mathbb{Z}_n$ and $\langle g \rangle$ respectively. Hence, by the theorem, Theorem 8.65, these images are isomorphic which is what we wanted to prove. ■

Notice in this case the induced map is $\mathbf{E}_g^n \colon \mathbb{Z}_n \to G$ satisfies

$$\mathbf{E}_g^n \circ \pi_n = \mathbf{E}_g$$

since, by definition, both sides take and integer m to g^m. (This is precisely the function $\mathbf{E}_g^d$ defined in Notation 5.60 back on page 281). Indeed, you can think of the proof of the above theorem as a generalization of the proof appearing in that corollary) Therefore, for a fixed integer n, no matter what cyclic group you come up with of order n, it must be isomorphic to $\mathbb{Z}_n$. For instance, each of the following six groups of order 6

$$\{\{6\}\} \qquad \mathbb{Z}_2 \times \mathbb{Z}_3 \qquad \mathcal{U}(\mathbb{Z}_7) \qquad \mathcal{U}(\mathbb{Z}_9) \qquad \mathcal{U}(\mathbb{Z}_{14}) \qquad \mathcal{U}(\mathbb{Z}_{18})$$

are cyclic. Hence each of them are isomorphic to $\mathbb{Z}_6$. Actually, we have

COROLLARY[2] 8.71.

(1) *For all prime numbers p, $\mathcal{U}(\mathbb{Z}_p) \cong \mathbb{Z}_{p-1}$*

(2) *If n and m are relatively prime, then $\mathbb{Z}_n \times \mathbb{Z}_m \cong \mathbb{Z}_{nm}$.*

(3) *For any integer n, the n^{th} roots of unity, $\{\{n\}\}$, are isomorphic to $\mathbb{Z}_n$.*

Proof. We have shown that each of the groups considered are cyclic of the correct order. ∎

Example 8.72. We have

$$G[\,n\,] \overset{1\text{-}1}{\leftrightsquigarrow} \hom(\mathbb{Z}_n, G).$$

That is $\mathbf{O}_n(G)$ (the cardinality of the LHS) is equal to the number of homomorphisms from $\mathbb{Z}_n$ to the group G (the cardinality of the RHS). This is because for all $g \in G$, the function $\mathbf{E}_g : \mathbb{Z} \to G$ from 16 back on page 391 has kernel equal to $|\,g\,|\mathbb{Z} \preccurlyeq \mathbb{Z}$. Meanwhile, the kernel of the map $\theta_n : \mathbb{Z} \to \mathbb{Z}_n$ is $n\mathbb{Z}$. Therefore, we have an induced map $\mathbb{Z}_n \to G$ iff $n\mathbb{Z} \subseteq |\,g\,|\mathbb{Z}$. But this happens iff $|\,g\,| \mid n$ which happens iff $g^n = e_G$.

The following exercise shows us that this theorem requires that the two kernels be literally the same, not just isomorphic!

Exercise 8.54. Show that the two homomorphisms

$$\begin{aligned} &: \mathbb{Z}_4 \times \mathbb{Z}_2 \to \mathbb{Z}_4 \\ &: ([\,a\,]_4, [\,b\,]_2) \mapsto [\,a\,]_4 \end{aligned} \qquad \begin{aligned} &: \mathbb{Z}_4 \times \mathbb{Z}_2 \to \mathbb{Z}_2 \times \mathbb{Z}_2 \\ &: ([\,a\,]_4, [\,b\,]_2) \mapsto ([\,a\,]_2, [\,b\,]_2) \end{aligned}$$

have isomorphic kernels, but they do not have isomorphic images.

Solution

Before the next exercise we first give a definition:

Definition 8.73. Two subgroups, H_1 and H_2, of a group G are *weakly conjugate* if there exists an automorphism of the group G which takes H_1 to H_2.

Notice that conjugate subgroups are weakly conjugate (since conjugation is an automorphism) and weakly conjugate subgroups are isomorphic. Thus we could just as easily call weakly conjugate groups, *strongly isomorphic* in G. I guess that would make conjugate subgroups really strongly isomorphic.

H and K Conjuate $\Rightarrow$ H and K Weakly Conjugate $\Rightarrow$ H and K isomorphic.

Example 8.74. The groups $\langle\,\mathbf{r}\,\rangle$ and $\langle\,\mathbf{r}\rho\,\rangle$ are weakly conjugate in $\mathcal{D}_4$ since the automorphism from Exercise 8.3 takes the first, $\langle\,\mathbf{r}\,\rangle$, to the second, $\langle\,\mathbf{r}\,\rangle$.

The next exercise shows despite the previous exercise, we can strengthen[VI] the theorem a bit.

Exercise 8.55. Suppose that two homomorphisms have weakly conjugate kernels. Prove that they have isomorphic images.

Solution

[VI] In other words we can find more situations where it applies than the stated case.

EXERCISE 8.56. Prove that the subgroups $\langle (2,0) \rangle$ and $\langle (0,1) \rangle$ are isomorphic to each other, but are not weakly conjugate subgroups of $\mathbb{Z}_4 \times \mathbb{Z}_2$.
Hint! Simply put the previous few exercises together.

SOLUTION

8.7 DIRECT PRODUCTS

Let H and K be subgroups of a group G. Define a function Define the map

$$\varphi \colon H \times K \to G \tag{7}$$
$$: (h,k) \mapsto hk. \tag{8}$$

In general, φ needn't be a homomorphism. However there is an important special case when it is a homomorphism.

THEOREM 8.75. *Suppose that*

$$\textit{every element of } H \textit{ commutes with every element of } K. \tag{9}$$

Then, the map φ defined in (7) is a homomorphism of groups which has kernel with $|H \cap K|$ elements and image HK.

PROOF. Let's break down the three claims.

(1) φ is a homomorphism.

Let $(h,k),(h',k') \in H \times K$. Then

$$\begin{aligned} \varphi((h,k)(h',k')) &= \varphi((hh',kk')) \\ &= \varphi((hh',kk')) \\ &= hh'kk' \\ &= hkh'k' \qquad \text{(By 9)} \\ &= \varphi(h,k)\varphi(h',k'). \end{aligned}$$

and so φ is operation preserving.

(2) $\operatorname{Ker}(\varphi)$ has $|H \cap K|$ elements in it.

I show this by proving that

$$\operatorname{Ker}(\varphi) = \{(h,h^{-1}) : h \in H \cap K\}$$

which clearly has $|H \cap K|$ elements in it.

$$\begin{aligned} \operatorname{Ker}(\varphi) &= \{(h,k) : h \in H, k \in K, \varphi((h,k)) = e_G\} \\ &= \{(h,k) : h \in H, k \in K, hk = e_G\} \\ &= \{(h,k) : h \in H, k \in K, h = k^{-1}\} \\ &= \{(h,h^{-1}) : h \in H \cap K\}. \end{aligned}$$

(3) $\mathrm{Im}(\varphi) = HK$.

This is obvious since by definition, $g \in HK$ means there exists $h \in H$ and $k \in K$ so that $g = hk$ which so happens to equal$\varphi((h,k))$. ∎

We will apply the above to a family of subgroups $H_1, \ldots, H_n$ where each element of H_i commutes with each element in H_j. This always holds for instance in an abelian group. The same basic proof will show that the corresponding map is still a homomorphism under this hypothesis, and moreover this homomorphism can just as easily seen to have image given by

$$H_1 H_2 \cdots H_n.$$

However, unfortunately, in this more general case we cannot as easily pick out how many elements are in the kernel of our homomorphism.

Exercise 8.57. Prove a converse of the above theorem: prove that if the function φ defined in (9) is a homomorphism, then every element of H commutes with every element of K.

SOLUTION

We can use the following lemma to make the requirement that every element of H commute with every element of K seem more natural.

Lemma 8.76. *Suppose that $H \trianglelefteq G$ and $K \trianglelefteq G$ and $H \cap K = \{e_G\}$. Then every element of H commutes with every element of K.*

Proof. Let $h \in H$ and $k \in K$ be arbitrary. Consider the element $w = hkh^{-1}k^{-1}$. Obviously, if we can show that this element $w = e_G$, we can conclude that h and k commute. We do this by showing that w is in $H \cap K$ which only contains the identity by the lemma's hypothesis. Now since $K \trianglelefteq G$ we know that $k' := hkh^{-1} \in K$. Therefore we have that

$$w = hkh^{-1}k^{-1} = k'k^{-1} \in K$$

by closure of the subgroup K under inverses and multiplication. Similarly, if we use the fact that $H \trianglelefteq G$ we can conclude that $w \in H$ as well. Hence $w \in H \cap K$. But our assumption is that only thing is in this intersection, namely the identity of G, e_G. Thus $w = e_G$, and as already pointed out this proves that h and k commute. Since these h and k were arbitrary the proof of the lemma is complete. ∎

Theorem 8.77 (Direct Product Recognition Theorem)**.** *Suppose $H, K \leqslant G$. Then G is isomorphic to the direct product of H and K, $H \times K$, iff*

DP 1 $H \cap K = \{e_G\}$;

DP 2 $HK = G$;

DP 3 $H, K \trianglelefteq G$.

Proof. Clearly the natural copies of H and K inside of $H \times K$ satisfy the three conditions, so we only need to prove the reverse direction.

That is suppose that H and K are subgroups of a group G which satisfy **DP 1**, **DP 2**, and **DP 3**. We need to prove that $G \cong H \times K$.

By Lemma 8.76, we know that every conditions **DP 1** and **DP 3** imply that every element of H commutes with every element of K. Therefore, we have access to homomorphism from Theorem 8.75.

One of the conclusions of that theorem is that the kernel of the map φ has $|H \cap K|$ elements in it. But the hypothesis, **DP 1**, says that this intersection is trivial. So the homomorphism φ is has trivial kernel, and, as such, must be injective.

The other conclusion from this theorem is that the image of the map φ is all of HK. But **DP 2** says that $HK = G$. As such under our assumptions, φ is also surjective.

Therefore under our hypothesis we have proven that φ is a bijective homomorphism, $\varphi\colon H \times K \to G$, i.e. an isomorphism. ∎

EXERCISE 8.58. Let G, G_1, G_2, A, A_1, and A_2, be groups with A, A_1, and A_2 abelian (written additively).

SOLUTION

(1) Prove that $\hom(G, A)$ forms an abelian group under the operation of function addition, i.e.

$$(\varphi + \psi)(g) := \varphi(g) + \psi(g).$$

(2) Prove that

$$\hom(G_1 \times G_2, A) \cong \hom(G_1, A) \times \hom(G_2, A)$$
$$\text{and}$$
$$\hom(G, A_1 \times A_2) \cong \hom(G, A_1) \times \hom(G, A_2).$$

Hint! Given homomorphisms $\varphi\colon G_1 \to A$ and $\psi\colon G_2 \to A$ define

$$\Omega\colon \hom(G_1, A) \times \hom(G_2, A) \to \hom(G_1 \times G_2, A)$$

via the rule

$$\Omega(\varphi, \psi)\colon G_1 \times G_2 \to A$$
$$: (g_1, g_2) \mapsto \varphi(g_1) + \varphi(g_2).$$

And define

$$\Psi\colon \hom(G_1 \times G_2, A) \to \hom(G_1, A) \times \hom(G_2, A)$$

via the rule

$$\Psi(\alpha) := (\alpha|_{G_1}, \alpha|_{G_2}).$$

Prove that either Ω or Ψ is operation preserving and that they are inverses to each other.

EXERCISE$^{\text{Hard!}}$ **8.59.** Prove that

$$\hom(\mathbb{Z}_n, \mathbb{Z}_m) \cong \mathbb{Z}_{\gcd(n,m)}.$$

SOLUTION

8.8 OTHER

Exercise 8.60. Suppose that $(M, \star)$ and $(N, *)$ are monoids and that $\varphi\colon M \to N$ is an operation preserving bijection.[I] Prove that $m \in M$ is invertible in $(M, \star)$ iff $\varphi(m)$ is invertible in $(N, *)$. Conclude that φ restricts to an isomorphism of groups $\mathcal{U}(M) \to \mathcal{U}(N)$.

Solution

Careful! You may need to reprove several things! For instance, when proving that a group homomorphism $\varphi\colon G \to H$ takes the identity of G to the identity of H we explicitly used the fact that H was a group![II]

Exercise 8.61. Suppose that n and m are positive integers.

Solution

(1) Prove that the map

$$\begin{aligned}\psi\colon \mathbb{Z}_{nm} &\to \mathbb{Z}_n \times \mathbb{Z}_m \\ : [\,a\,]_{n\cdot m} &\mapsto ([\,a\,]_n\,, [\,a\,]_m)\end{aligned}$$

preserves both addition and multiplication.[III]

Now we assume that n and m are relatively prime.

(2) Prove that, in this case, ψ is a bijection.

Hint! Since both sets are finite and have the same order, one simply needs to show that ψ is injective. To do this you may care to use the fact that the lcm of two relatively prime integers is their product.

(3) Use the previous exercise to prove, in this case,

$$\mathbb{Z}_{nm} \cong \mathbb{Z}_n \times \mathbb{Z}_m \quad \text{and} \quad \mathcal{U}(\mathbb{Z}_{nm}) \cong \mathcal{U}(\mathbb{Z}_n) \times \mathcal{U}(\mathbb{Z}_m).$$

Based on this previous exercise, we have the following Corollary which finally completes the proof of the formula for Euler totient function which we originally presented back on page 153.

Corollary 8.78. *If n and m are relatively prime, then $\vartheta(n \cdot m) = \vartheta(n) \cdot \vartheta(m)$ where ϑ is the Euler totient function.*

Proof. From the exercise, Exercise 8.61, we know that whenever n and m are relatively prime

$$\begin{aligned}&\mathcal{U}(\mathbb{Z}_n) \times \mathcal{U}(\mathbb{Z}_m) \cong \mathcal{U}(\mathbb{Z}_{n\cdot m}) \\ &\Rightarrow |\,\mathcal{U}(\mathbb{Z}_n) \times \mathcal{U}(\mathbb{Z}_m)\,| = |\,\mathcal{U}(\mathbb{Z}_{n\cdot m})\,| && \text{(Iso. gps have same order)} \\ &\Rightarrow |\,\mathcal{U}(\mathbb{Z}_n)\,| \cdot |\,\mathcal{U}(\mathbb{Z}_m)\,| = |\,\mathcal{U}(\mathbb{Z}_{n\cdot m})\,| && \text{(For sets, } \#(A \times B) = \#A \cdot \#B) \\ &\Rightarrow \vartheta(n) \cdot \vartheta(m) = \vartheta(n \cdot m) && (\forall r \in \mathbb{Z}^+, \vartheta(r) \overset{\text{def}}{=} |\,\mathcal{U}(\mathbb{Z}_r)\,|).\blacksquare\end{aligned}$$

As pointed out at the time, in Remark 3.4.3, this corollary proves that for any integer n that

$$\vartheta(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right).$$

Let n be an integer which factors into primes as

$$n = p_1^{m_1} \cdots p_k^{n_k}.$$

Then, using the above, we have

$$\mathcal{U}(\mathbb{Z}_n) \cong \mathcal{U}(\mathbb{Z}_{p_1^{m_1}}) \times \cdots \times \mathcal{U}(\mathbb{Z}_{p_k^{m_k}}).$$

And from our general theory, we know that this is a cyclic group iff each factor is cyclic and the order of the factors are relatively prime to each other. Since for a prime p we have

$$\left|\mathcal{U}(\mathbb{Z}_{p^r})\right| = \vartheta(p^r) = p^{r-1} \cdot (p-1)$$

the only way that $\left|\mathcal{U}(\mathbb{Z}_{p^r})\right|$ is odd is if $r = 1$ and $p = 2$. Moreover, we have seen previously that

$$\mathcal{U}(\mathbb{Z}_{2^r})$$

is cyclic iff $r = 1$ or $r = 2$.

PROPOSITION 8.79. *If p is an odd prime, then for any positive integer n, $\mathcal{U}(\mathbb{Z}_{p^r})$ is a cyclic group.*

PROOF. Since

$$(1+p)^{p^{r-2}} \equiv_{p^r} 1 + p^{r-2}p = 1 + p^{r-1} \text{ and } (1+p)^{p^{r-1}} \equiv_{p^r} = 1$$

we have that in $\mathcal{U}(\mathbb{Z}_{p^r})$ the element $[\,1+p\,]_{p^r}$ has order p^{r-1}.

Meanwhile, noting that we have functions

$$\pi_s^r\colon\ [\,z\,]_{p^r} \mapsto [\,z\,]_{p^s}$$

whenever $s \leq r$ which preserve both addition and multiplication and as such induce homomorphisms of both the additive group. Now let's take an integer a in such a way that $[\,a\,]_p$ is generator of the cyclic group $\mathbb{Z}_p^\times$. This means that $[\,a\,]_p$ has order $p-1$. As such, the order of $[\,a\,]_{p^r}$ which maps to this generator under π_1^r, must be divisible by $(p-1)$. In particular, some power of $[\,a\,]_{p^r}$, call it y has order exactly $p-1$. Therefore, since p is odd, $p-1$ and p^{r-1} are relatively prime and so the element $[\,y \cdot (1+p)\,]_{pp^r}$ has order $p^{r-1} \cdot (p-1)$, which is the order of the group $\mathcal{U}(\mathbb{Z}_{p^r})$ itself. Therefore this group must be cyclic. ∎

Since $\mathcal{U}(\mathbb{Z}_2)$ is the trivial group, in this case $\mathcal{U}(\mathbb{Z}_{2p^r}) \cong \mathcal{U}(\mathbb{Z}_{p^r})$ and so $\mathcal{U}(\mathbb{Z}_{2p^r})$ is also cyclic. Lastly, we note that in Exercise 5.59 back on page 294 you show that the groups $\mathcal{U}(\mathbb{Z}_{2^r})$ are not cyclic for $r > 2$, i.e. for powers of 2 besides 2 and 4. Therefore we have proven

THEOREM 8.80. *The group $\mathcal{U}(\mathbb{Z}_n)$ is cyclic iff*

(1) *n is a power of an odd prime, or*

(2) *n is twice a power of an odd prime, or*

(3) *$n = 4$.*

In Proposition 13.63 we'll go on to classify the last remaining piece to the puzzle, namely the nature of $\mathcal{U}(\mathbb{Z}_{2^r})$. Indeed, we'll show that despite the fact that these groups are not cyclic, they come pretty close. Indeed, we prove that

$$\mathcal{U}(\mathbb{Z}_{2^r}) \cong \mathbb{Z}_2 \times \mathbb{Z}_{2^{r-2}}.$$

(Of course $\mathcal{U}(\mathbb{Z}_2)$ is the trivial group consisting only of the element $[\,1\,]_2$ by itself.)

The following exercise generalizes Theorem 5.113 back on page 310.

EXERCISE 8.62. Suppose n factors into primes as $p_1^{m_1} \cdots p_r^{m_r}$. Define

$$\mathcal{L} = \text{lcm}(\vartheta(p_1^{m_1}), \ldots, \vartheta(p_r^{m_r})).$$

Prove that for any integer a relatively prime with n that

$$a^{\mathcal{L}} \equiv_n 1.$$

Hint! First prove this for n a power of an odd prime, then use the isomorphism

$$: \mathcal{U}(\mathbb{Z}_n) \xrightarrow{\cong} \mathcal{U}(\mathbb{Z}_{p_1^{m_1}}) \times \cdots \times \mathcal{U}(\mathbb{Z}_{p_r^{m_r}})$$
$$: [a]_n \mapsto ([a]_{p_1^{m_1}}, \ldots, [a]_{p_r^{m_r}}).$$

SOLUTION

In particular, since $\vartheta(n)$ is a multiple of $\mathcal{L}$ from this exercise, we also have that $a^{\mathcal{L}} \equiv_n 1$ in this situation. We can use this a few times to simplify thing considerably. For instance:

EXAMPLE 8.81. Let's use the methods developed to find the first 2 digits of $7^{7^{100}}$.[IV] Well, the first two digits of a number is really the same thing as that number mod 100. So we want to calculate

$$7^{7^{100}} \bmod 100.$$

In terms of the previous exercise, since

$$100 = 5^2 \cdot 2^2 \Rightarrow \mathcal{L} = \text{lcm}\left(\vartheta(5^2), \vartheta(2^2)\right) = \text{lcm}\,(20, 2) = 20$$

and since 7 is relatively prime with 100, we know that $7^{20} \equiv_{100} 1$. As such we know that

$$7^{7^{100}} \equiv_{100} 7^{7^{100} \bmod 20}.$$

So we need to figure out what the exponent is mod 20:

$$7^{100} \bmod 20.$$

Well, again, 7 is relatively prime with 20 so we can apply the previous exercise again to aide us with this calculation. Now we have a new $\mathcal{L}$ which this time is given by

$$\text{lcm}\,(\vartheta(5), \vartheta(4)) = 4.$$

Moreover since $100 \bmod 4 = 0$, we know that

$$7^{100} \equiv_{20} 7^0 = 1.$$

Hence

$$7^{7^{100}} \equiv_{100} 7^{7^{100} \bmod 20} = 7^1 = 7.$$

In conclusion, the last 2 digits of $7^{7^{100}}$ are 0 and 7.

[IV] Computing this directly even in a mathematical software like MAPLE crashes it.

Since we can now write $\mathcal{U}(\mathbb{Z}_n)$ as a product of cyclic groups, we can now easily compute the order tables of these groups. For instance, if we take

$$n = 720 = 5 \cdot 3^2 \cdot 2^4$$

then

$$|\mathcal{U}(\mathbb{Z}_{720})| = \vartheta(720) = \vartheta(5)\vartheta(3^2)\vartheta(2^4) = 4 \cdot 6 \cdot 8 = 192.$$

More importantly

$$\begin{aligned}\mathcal{U}(\mathbb{Z}_{720}) &\cong \mathcal{U}(\mathbb{Z}_5) \times \mathcal{U}(\mathbb{Z}_9) \times \mathcal{U}(\mathbb{Z}_{16}) \\ &\cong \mathbb{Z}_4 \times \mathbb{Z}_6 \times (\mathbb{Z}_2 \times \mathbb{Z}_4) \\ &\cong \mathbb{Z}_4 \times \mathbb{Z}_6 \times \mathbb{Z}_2 \times \mathbb{Z}_4. \\ &\cong \mathbb{Z}_{12} \times \mathbb{Z}_4 \times \mathbb{Z}_2^2.\end{aligned}$$

As such, we have that every element in $\mathcal{U}(\mathbb{Z}_{720})$ has order dividing $\text{lcm}(4, 6, 2, 4) = 12$. We can also easily calculate the $\mathbf{O}_r$'s. For instance,

$$\mathbf{O}_6(\mathcal{U}(\mathbb{Z}_{720})) = \mathbf{O}_6(\mathbb{Z}_4 \times \mathbb{Z}_6 \times \mathbb{Z}_2 \times \mathbb{Z}_4) = \mathbf{O}_6(\mathbb{Z}_4) \cdot \mathbf{O}_6(\mathbb{Z}_6) \cdot \mathbf{O}_6(\mathbb{Z}_2) \cdot \mathbf{O}_6(\mathbb{Z}_4) = 2 \cdot 6 \cdot 2 \cdot 4 = 48.$$

Making a table

n	1	2	3	4	6	12
$\mathbf{O}_n$	1	16	3	64	48	192

.

Now we finish this off by using Theorem 5.92 to calculate the order table:

$$\begin{aligned}\mathbf{o}_{12} &= \mathbf{O}_{12} - (\mathbf{O}_6 + \mathbf{O}_4) + \mathbf{O}_2 = 192 - (48 + 64) + 16 = 96 \\ \mathbf{o}_6 &= \mathbf{O}_6 - (\mathbf{O}_3 + \mathbf{O}_2) + \mathbf{O}_1 = 48 - (3 + 16) + 1 = 30 \\ \mathbf{o}_4 &= \mathbf{O}_4 - \mathbf{O}_2 = 64 - 16 = 2^6 - 2^4 = 48 \\ \mathbf{o}_3 &= \mathbf{O}_3 - \mathbf{O}_1 = 3 - 1 = 2 \\ \mathbf{o}_2 &= \mathbf{O}_2 - \mathbf{O}_1 = 16 - 1 = 15 \\ \mathbf{o}_1 &= \mathbf{O}_1 = 1.\end{aligned}$$

which can be organized into the order table:

n	1	2	3	4	6	12
$\mathbf{o}_n$	1	15	2	48	30	96

.

Definition 8.82. Let G and H be groups and $\mathcal{R}$: $G \rightsquigarrow H$ a relation. We say that $\mathcal{R}$ is *operation preserving* if

$$g_1 \mathcal{R} h_1 \text{ and } g_2 \mathcal{R} h_2 \Rightarrow (g_1 g_2) \mathcal{R} (h_1 h_2)$$

and *inverse preserving* if

$$g \mathcal{R} h \Rightarrow g^{-1} \mathcal{R} h^{-1}.$$

We say that a relation $\mathcal{R} \subseteq G \rightsquigarrow H$ is a *hom relation* if it is both operation preserving and inverse preserving.

Notice that saying that $\mathcal{R}$ is operation preserving is the same thing as saying that it is closed under multiplication.

Exercise 8.63. Suppose that $\mathcal{R}\colon G \rightsquigarrow H$ and $\mathcal{S}\colon H \rightsquigarrow K$ are hom relations.

Solution

(1) Prove that $\mathcal{S} \circ \mathcal{R}$ and

(2) $\mathcal{R}^{-1}$

are also hom relations.

Exercise 8.64.

Solution

(1) Prove that (the graph[V] of) any homomorphism $G \to H$ is a hom relation.

(2) Suppose that a hom relation is (the graph of) a function (i.e. it is defined everywhere and single valued). Prove that it is (the graph of) a group homomorphism.

Exercise 8.65.

Solution

(1) Prove that any subgroup of $G \times H$ is a hom relation $G \rightsquigarrow H$.

(2) Conversely, prove that any non-empty hom relation $\mathcal{R}\colon G \rightsquigarrow H$ is a subgroup of $G \times H$.

Exercise 8.66. Let $\mathcal{R}\colon G \rightsquigarrow H$ be a o.p. which is defined everywhere and assume that H is finite.

Solution

(1) Prove that

(a) $e_G \mathcal{R} e_H$

(b) $e_G \mathcal{R} h \Rightarrow e_G \mathcal{R} h^{-1}$.

Hint! First prove $e_G \mathcal{R} h \Rightarrow e_g \mathcal{R} h^n$ for any positive integer n, then use the fact that every element of a finite group has finite order.

(2) Prove that

$$g \mathcal{R} h \Leftrightarrow g^{-1} \mathcal{R} h^{-1}.$$

Hint! Using fact $\mathcal{R}$ is defined everywhere all we know is that $g \mathcal{R} h$ and $g^{-1} \mathcal{R} h'$ to some $h' \in H$. Use the above two exercises to prove that $e_G \mathcal{R} h'^{-1}h^{-1}$ and use that to conclude that $g^{-1} \mathcal{R} h^{-1}$.

(3) Conclude that $\mathcal{R}$ is a hom relation.

Definition 8.83. Given a hom relation $\mathcal{R}\colon G \rightsquigarrow H$ define

$$\mathrm{Ker}\,(\mathcal{R}) \stackrel{\mathrm{def}}{=} \mathcal{R}^{-1}(e_H) \quad \text{and} \quad \mathrm{Im}\,(\mathcal{R}) \text{ as always.}$$

The following exercise can easily be adapted to another proof of Lemma 8.66.

Exercise 8.67.

(1) Prove that a hom relation $\mathcal{R}$ is single valued iff $\mathrm{Ker}\left(\mathcal{R}^{-1}\right) = \{e_G\}$.

(2) Prove that a hom relation $\mathcal{R}$ is defined everywhere iff $\mathrm{Im}\left(\mathcal{R}^{-1}\right) = G$.

(3) State a condition which guarantees us that $\mathcal{R}$ is (the graph of) a group homomorphism

$$\varphi\colon G \to H.$$

SOLUTION

9 Generators

Generators are the dream of every lazy group theorist. They give us a way of saving a lot of writing by simply talking about a small subset of a group. This is akin the red, yellow, cyan scheme for colors. Clearly if you open up your printer, you do not set see tens of thousands of colors that your printer is capable of printing, but only the three mentioned. In some sense, those three basic colors *generate* all other colors in the sense that mixing those three colors together gives you most other colors. In group theory, we don't mix; we multiply. So a rough definition of what it means for a subset X of a group G to generate G is that every element of G can be written as a products of elements of X.

9.1 Basic Definitions

We restate Definition 6.18 back on page 318.

Definition 9.1. For any subset $X \subseteq G$ of a group. Then $\langle X \rangle$ is defined as the intersection of all subgroups of G which contain X. It is called the group *generated by* X. If a group $H = \langle X \rangle$ for some set, we say that H is *generated* by X.[I]

We therefore have two natural questions:

(1) Given a group what is a generating set for that group?[II]

(2) Given a subset of a group what is the group it generates?

Since G is a group containing X, we are intersecting over a non-empty collection of subgroups each containing X. Hence $\langle X \rangle$ is indeed a group by Lemma 6.17 and it contains X by Proposition 2.22. This same proposition also tells us that it is contained inside of any group which contains X. That is $\langle X \rangle$ has the following three properties (which should be thought of as saying that $\langle X \rangle$ is the *smallest* subgroup of G which contains X):

Suppose that $X \subseteq G$. Then $\langle X \rangle$ is the smallest subgroup of G which contains X. That is $\langle X \rangle$ satisfies the following three properties:

gen 1 $X \subseteq \langle X \rangle$;

gen 2 $\langle X \rangle \leqslant G$;

gen 3 If $H \leqslant G$, then $X \subseteq H \Rightarrow \langle X \rangle \leqslant H$.

[I] It might appear that $\langle X \rangle$ depends on G, but we shall see that it does not.

[II] For any group G, the set G generates G! However, if we take this as our generating set we are not being that lazy! When finding a generating set we typically want to take the generating set to be as small as possible!!

The following proposition tells us that we could use these three properties as the definition of $\langle X \rangle$. That is, we could have made the following definition:

Given X, $\langle X \rangle$ is the unique group which satisfies **gen** 1, **gen** 2, and **gen** 3 above.

PROPOSITION 9.2. *If K is a group which satisfies* **gen** 1, **gen** 2, *and* **gen** 3 *above (for a given X), then $K = \langle X \rangle$.*

PROOF. If two groups have these three properties, then by **gen** 3 they are each contained inside of each other, so they are indeed equal. ∎

REMARK 9.3.

(1) For any group H, we have $H = \langle H \rangle$. In fact H is a group iff $H = \langle H \rangle$.

(2) $\langle \varnothing \rangle = \{e\}$. Indeed, in this very special case, $\{e\}$ trivially satisfies the three properties.

(3) If $X \subseteq Y$, then since $X \subseteq Y \subseteq \langle Y \rangle$ and $\langle Y \rangle$ is a group, we have by **gen** 3 that $\langle X \rangle \subseteq \langle Y \rangle$.

(4) If either H normalizes K or K normalizes H, then we have by Theorem 6.48 that $\langle H \cup K \rangle = HK$. However, in general, we always have that $HK \subseteq \langle H \cup K \rangle$

DEFINITION 9.4. Let X be a set contained inside of some groups G. We define the words in X, denoted $W(X)$ inductively as follows:

- The identity, $e \in W(X)$.
- If $w \in W(X)$ and $x \in X$, then both wx and $wx^{-1} \in W(X)$.

Thus elements of $W(X)$ are simply elements of the form

$$x_1^{n_1} \cdots x_m^{n_m}$$

where the n_i are allowed to be any integer and the x_i are each elements of X.

WARNING! 9.5. Taking the definition of words to be the formula is very misleading and has led numerous students off of HUGE cliffs! The subscripts on the x's only represent the state of the induction in which the x was tacked on. For instance we could think of x^4 as taking 1 stage, 4 stages or even more (if we were dumb enough to say start with x^6 and then started tacking on inverses).

Except for the presence of negative powers this is exactly what we mean when we talk about words in the usual alphabet of letters (except the words in group theory don't actually have to spell anything). In particular every single word that appears in this book is an example of a word in the usual latin alphabet, (except for the word $\kappa o \ell \ell$ which I just made up ...). Since powers of x are just products of x or x^{-1} with itself, we could have written the definition by taking the integer n to be either 1 or -1. Notice that $X \subseteq W(X)$ since every $x \in X$, stage 1 of the inductive definition says that $e \in W(X)$, and stage 2 says that $x = e \cdot x \in W(X)$.

EXAMPLE 9.6.

(1) If $X = \{a, b\}$, then the elements a, b, b^2a, aba, aa^{-1} and $a^3bab^{-1}a^{-5}$ are all words in X. Notice that the product of two words is also a word. For instance aba times $a^3bab^{-1}a^{-5}$ is simply

$$(aba)(a^3bab^{-1}a^{-5}) = aba^4bab^{-1}a^{-5}.$$

(2) If X consists of a single element a, then $W(\{a\})$ is simply the set

$$\{a^i : i \in \mathbb{Z}\}$$

of powers of a, which was previously denoted as $\langle a \rangle$ (foreshadowing??) (you can't spell a lot of words with 1 letter!).

THEOREM 9.7. $W(X) = \langle X \rangle$.

PROOF.

$\subseteq$ Since $\langle X \rangle$ is a group, we have $e \in \langle X \rangle$ and so it contains the first stage of the induction. Moreover $\langle X \rangle$ also contains x and x^{-1} for every $x \in X$ (the former by definition, the latter by closure of the group $\langle X \rangle$ under taking inverses). Therefore, if $w \in \langle X \rangle$, then by closure we also have that $wx \in \langle X \rangle$ and $wx^{-1} \in \langle X \rangle$. Therefore $W(X) \subseteq \langle X \rangle$.

$\supseteq$ Since $W(X)$ contains X, we only need to show that it is a subgroup of G and apply **gen** 3 to prove the reverse inclusion: $\langle X \rangle \subseteq W(X)$.

- **Closure** As we already pointed out, the product of two words is a word.
- **Inverses** The inverse of a word is a word by generalized socks and shoes, Exercise 4.15. ■

COROLLARY 9.8. $\langle Y \rangle \leqslant \langle X \rangle$ *iff every element in* Y *can be written as a word in* X. ■

Using this corollary in two two directions we have that

$\langle X \rangle = \langle Y \rangle$ iff every letter in X can be written as a word in Y and every letter in Y can be written as a word in X.

REMARK 9.9. If you've had linear algebra before this should look familiar in terms of a basis of vector space. Indeed, the span of a set of vectors in a vector space is the "smallest" vector subspace containing the set of vectors in exactly the same way that the group generated by a set of elements in the "smallest" subgroup containing that set. However, as is almost always true, things are simpler in the linear algebra world since everything is commutative there.

EXAMPLE 9.10. In more detail, the above corollary means that if

$$X = \{x_1, \ldots, x_n\} \quad \text{and} \quad Y = \{y_1, \ldots, y_m\}$$

and for each y_i, there exists a word $y_i = w_i(x_1, \ldots, x_n)$, then $\langle Y \rangle \subseteq \langle X \rangle$.[III] You can think of this as solving a sequence of equations

$$\begin{aligned} y_1 &= w_1(x_1, \ldots, x_n) \\ y_2 &= w_2(x_1, \ldots, x_n) \\ &\vdots \quad \vdots \\ y_m &= w_m(x_1, \ldots, x_n). \end{aligned}$$

[III] Of course assuming the sets X and Y are finite is done only for notational concreteness.

This gives us a ton of ways of writing different generators for the same subgroup. For concreteness, let's stick with looking at a group generated by a two element set $X := \{x_1, x_2\}$. Then, if define $y_1 := x_1x_2$, $y_2 := x_2$, then by definition, both y_1 and y_2 are words in x_1 and x_2. Therefore, $\langle Y \rangle \subseteq \langle X \rangle$. On the other hand,

$$x_1 = y_1y_2^{-1} \quad \text{and} \quad x_2 = y_2$$

proves that we also have that $\langle X \rangle \subseteq \langle Y \rangle$ and so $\langle X \rangle = \langle Y \rangle$. For a more complicated example, consider

$$z_1 := x_2x_1x_2 \quad \text{and} \quad z_2 := x_2^{-1}x_1^{-1}x_2x_1x_2.$$

Then we have that

$$\begin{aligned}
z_1z_2z_1^{-1} &= (x_2x_1x_2)\left(x_2^{-1}x_1^{-1}x_2x_1x_2\right)(x_2x_1x_2)^{-1} \\
&= (x_2x_1x_2)\left(x_2^{-1}x_1^{-1}x_2x_1x_2\right)\left(x_2^{-1}x_1^{-1}x_2^{-1}\right) \\
&= x_2x_1(\cancel{x_2x_2^{-1}})x_1^{-1}x_2x_1(\cancel{x_2x_2^{-1}})x_1^{-1}x_2^{-1} \\
&= x_2x_1x_1^{-1}x_2x_1x_1^{-1}x_2^{-1} \\
&= x_2(\cancel{x_1x_1^{-1}})x_2(\cancel{x_1x_1^{-1}})x_2^{-1} \\
&= x_2x_2x_2^{-1} \\
&= x_2(\cancel{x_2x_2^{-1}}) \\
&= x_2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
z_1z_2^{-1}z_1z_2^{-1}z_1^{-1} &= \left(z_1z_2z_1^{-1}\right)^{-1} z_1 \left(z_1z_2z_1^{-1}\right)^{-1} \\
&= x_2^{-1}x_2x_1x_2x_2^{-1} \\
&= x_1.
\end{aligned}$$

REMARK 9.11. By this theorem, we see that $\langle X \rangle = \{x^n \mid n \in \mathbb{Z}\}$ (as already pointed out in Example 9.6(**2**) these are the only words one can spell with a single letter!).

9.2 EXAMPLES

9.2.1 GENERAL EXAMPLES

EXAMPLE 9.12.

(1) The group $\mathbb{Z}$ (under **ADDITION** OF COURSE!!!!!) is generated by 1. We already pointed this out in Remark 9.11.

(2) We also already pointed out that the groups $\mathbb{Z}_n$ are each generated by $[\,1\,]_n$. Again, any subset of $\mathbb{Z}_n$ closed under addition and containing $[\,1\,]_n$ must contain all of

$$\mathbb{Z}_n = \{[\,0\,]_n, [\,1\,]_n, [\,2\,]_n, \ldots, [\,n-1\,]_n.\}$$

(3) We have seen that every element of $\mathcal{D}_n$ can be written as a product of powers of $\varrho_{\frac{2\pi}{n}}$ (which we agreed to always call ρ) and powers of some fixed reflection $\mathbf{r}$ (which if we think of a polygon as a subset of $\mathbb{C}$ we have been taken as the reflection determined by complex conjugation). In other words $\mathcal{D}_n$ is generated by these two symmetries.

(4) Every element of $B\mathcal{D}_n$ is, by definition, a product of $\chi_{2n} = e^{\frac{\pi \mathbf{i}}{n}}$ and j. Thus these two quaternions generate this group. In fact, a more solid definition of $B\mathcal{D}_n$ could have been:

$B\mathcal{D}_n$ is the subgroup of quaternions generated by $\mathbf{j}$ and χ_{2n}.

(5) Consider the set of all positions of the $2 \times 2 \times 2$ Rubik's cube, $\mathcal{RK}_2$. It's better to think of the elements as the motion necessary to get from the starting position to a given configuration. Since every move gives a permutation of the $6 \cdot 4 = 24$ faces of the cube, we can think of this as a group which, after picking a labeling of the faces, is a subgroup of $\mathcal{S}_{224}$. We can generate this group by rotating the 6 faces by 90 degrees. This is actually quite redundant. For instance, we can rotate the back face by first rotating the two side faces 180 degrees, then rotating the front face, and finally by again rotating the two side faces by 180 degrees. Doing this sort of thing we can see that we can generate this group by rotating by 90 degrees along the front face, the left face, and the right face. This, however, is one example of a minimal generating set which is not very practical. It would be a pain in the butt if every time we wanted to rotate the back face we had to do a sequence of all of these moves. If instead we were dealing with the motions of a $3 \times 3 \times 3$ Rubik's cube, $\mathcal{RK}_3$, we have $9 \cdot 6 = 54$ (9 faces per side and 6 sides on a cube) faces, and we can think of the group in question as a subgroup of $\mathcal{S}_{54}$. If we understand these groups well enough, it is possible to easily solve any Rubik's cube puzzle.

9.2.2 Examples pertaining to $\mathcal{S}_n$.

Example 9.13.

(1) Every element of $\mathcal{S}_n$ can be written as a product of disjoint cycles. In particular every element of $\mathcal{S}_n$ can be written as a product of cycles and so cycles generate $\mathcal{S}_n$.[I]

(2) Improving on the previous example, I claim that $\mathcal{S}_n$ is generated by transpositions (which are $\underline{2}$ –cycles ... thus we don't need all cycles, but only $\underline{2}$–cycles). We only need to show that every cycle can be written as a product of transpositions. But cut and paste says exactly this for

$$(a_1 \cdots a_m) = (a_1\ a_2)(a_2\ a_3)\cdots(a_{m-1}\ a_m).$$

is a way of writing an $\underline{m}$–cycle as a product of transpositions.

(3) Improving on this even further I claim that $\mathcal{S}_n$ is generated by the $n-1$ *adjacent* transpositions $\{(1\ 2),(2\ 3),\ldots,(n-1\ n)\}$. Now we only need to show that every transposition $(i\ j)$ can be written as a product of these. Indeed, assuming that $i < j$ we have that

$$x_0 := (i\ i+1), x_1 := (i+1\ i+2), \ldots, x_{j-1} := (j-1\ j)$$

are all adjacent transpositions and by $\mathcal{S}_n$'s conjugation trick (Theorem 7.38**(2)**)

$$x_{j-1}(i\ j)x_{j-1}^{-1} = (i\ j-1)$$

and inductively

$$\left(x_1 \cdots x_{j-1}\right)(i\ j)\left(x_1 \cdots x_{j-1}\right)^{-1} = (i\ i+1)$$

[I] This generating set kinda blows ... for instance every element of $\mathcal{S}_3$ is a cycle and all but 3 elements of $\mathcal{S}_4$ are cycles. There should be a better generating set than this crap.

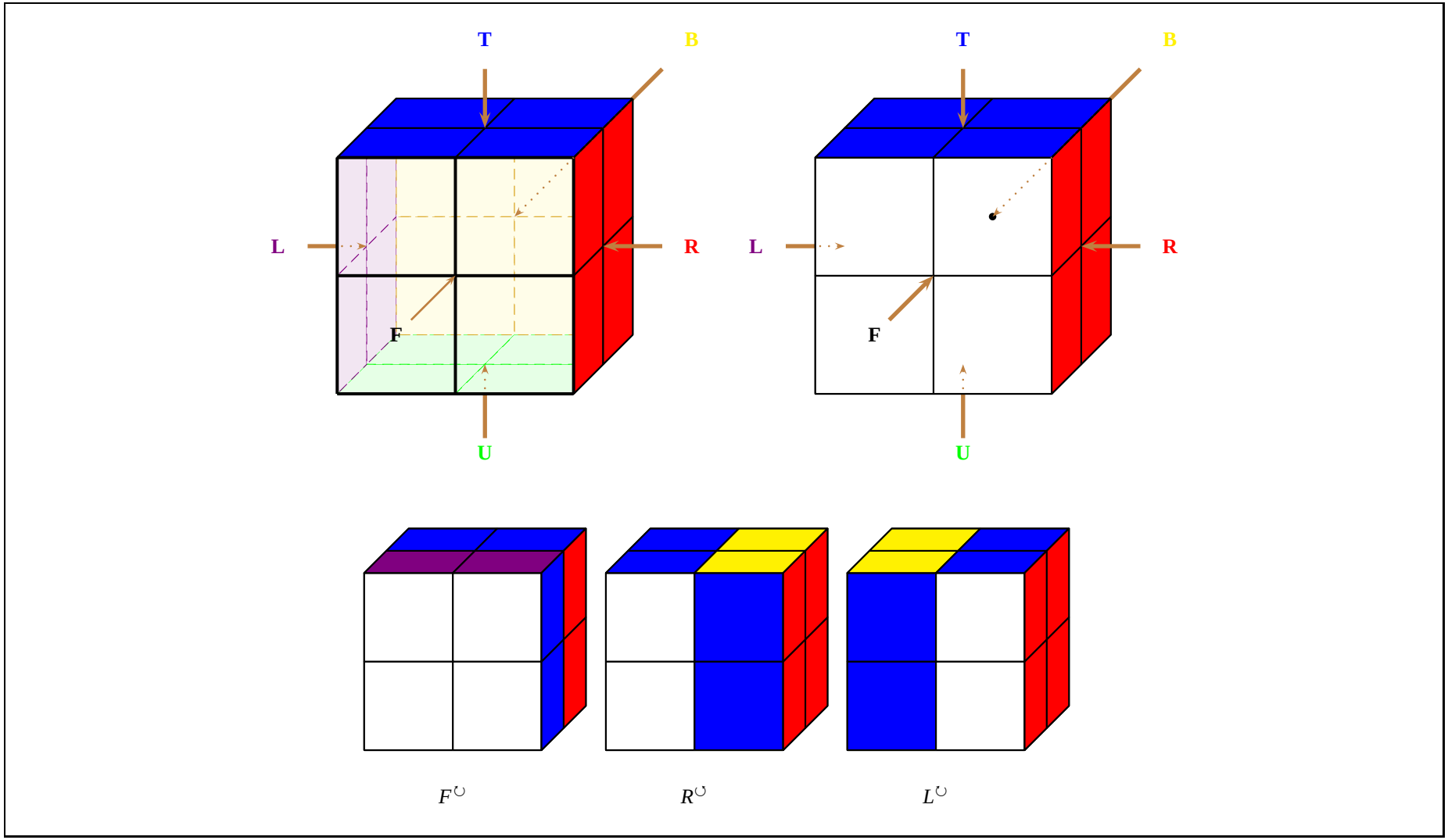


Figure 9.1: A $2 \times 2 \times 2$ Rubik's cube along with (the end result of applying) three basic moves. The faces are labeled F=Front, L=Left, R=Right, B=Back (not bottom!), U=Under, T=Top.

which is an adjacent transposition ($= x_0$). Hence we have that

$$\begin{aligned}\left(x_{j-1}\cdots x_1\right)(i\ i+1)\left(x_{j-1}\cdots x_1\right)^{-1} &= (i\ j)\\ \Rightarrow \left(x_{j-1}\cdots x_1\right)(i\ i+1)\left(x_1\cdots x_{j-1}\right) &= (i\ j)\end{aligned}$$

which shows us that the arbitrary transposition $(i\ j)$ can be written as a product of adjacent transpositions.

(4) We give yet another generating set for $\mathcal{S}_n$; this time using only two generators. We show that $\mathcal{S}_n$ can always be generated by the two elements $y := (1\ 2)$ and $x := (1\ 2\ \ldots\ n)$. Since we have just shown that adjacent transpositions generate $\mathcal{S}_n$, our goal is to show that every adjacent transposition $(i\ i+i)$ can be written as a product of x's and y's. But this an easy induction (using $\mathcal{S}_n's$ conjugacy trick again) that x^iyx^{-i} is equal to just that, so we are done. The following table demonstrates this explicitly in the case $n = 4$ in the minimum possible ways. Notice that this table isn't really all that nice. In particular many elements must be written as products that alternate between powers of x and y several times (which surprises many first year group theorists!)

$$\begin{aligned}(1) &= (x^4 = y^2)\\ (1\ 2) &= y\\ (1\ 3) &= yxyx^{-1}y = x^{-1}yx^2y = (yxyx^3y)\\ (1\ 4) &= x^{-1}yx = (yxyx^2)\\ (2\ 3) &= xyx^{-1} = (xyx^3)\\ (2\ 4) &= xyx^2y = (yx^2yx^{-1})\\ (3\ 4) &= x^2yx^2\\ (1\ 2\ 3) &= x^{-1}yx^2 = yxyx^{-1} = (yx^2yxy)\\ (1\ 2\ 4) &= x^2yx^{-1}\\ (1\ 3\ 2) &= x^2yx = xyx^{-1}y\\ (1\ 3\ 4) &= xy\\ (1\ 4\ 2) &= xyx^2 = yx^{-1}yx\end{aligned}$$

$$\begin{aligned}(1\ 4\ 3) &= yx^{-1} = (yx^3 = xyxy)\\ (2\ 3\ 4) &= yx\\ (2\ 4\ 3) &= x^{-1}y = (x^3y)\\ (1\ 2\ 3\ 4) &= x\\ (1\ 2\ 4\ 3) &= xyx = yx^{-1}y\\ (1\ 3\ 2\ 4) &= yx^2\\ (1\ 3\ 4\ 2) &= yxy = x^{-1}yx^{-1}\\ (1\ 4\ 2\ 3) &= x^2y\\ (1\ 4\ 3\ 2) &= x^{-1} = (x^3)\\ (1\ 2)(3\ 4) &= x^2yx^2y = xyx^{-1}yxy\\ (1\ 3)(2\ 4) &= x^2\\ (1\ 4)(2\ 3) &= yx^2y = (xyx^2yx)\end{aligned}$$

Exercise 9.1. Show the crossing number of a permutation σ is 1 iff σ is an adjacent transposition.

SOLUTION

Exercise 9.2. We generalize the above exercise. Show that the crossing number of a permutation σ is equal to the minimal number of adjacent permutations whose product is equal to σ.

Hint! (**This problem is a bit hard!**) Show that for any permutation σ there exists an adjacent transposition τ so that $\tau\sigma$ has crossing number one less than the crossing number of σ. Then proceed by induction to show there are adjacent transpositions $\tau_1, \ldots, \tau_{n-1}$ (where n is the crossing number of σ) so that

$$\tau_{n-1}\cdots\tau_1\sigma$$

has crossing number one.

SOLUTION

EXERCISE 9.3. Prove that the transpositions

$$(1\,2),\ (1\,3),\ \ldots,\ (1\,n)$$

generate $\mathcal{S}_n$.

SOLUTION

EXERCISE 9.4. Prove that

$$(1\,2), \quad \text{and} \quad (2\,3\cdots n)$$

generate $\mathcal{S}_n$.

SOLUTION

The next one is a bit harder.

EXERCISE 9.5. Prove that $\mathcal{S}_7$ is generated by

$$(1\,2\,3\,4) \quad \text{and} \quad (3\,4\,5\,6\,7).$$

SOLUTION

Hint! I don't have any great advice. Just play with it long enough until you get a $(1\,2\,3\,4\,5\,6\,7)$ and $(1\,2)$ to fall out and use the fact that these two elements generate $\mathcal{S}_7$.

EXAMPLE / THEOREM 9.14. Suppose that $N \trianglelefteq \mathcal{S}_n$ contains a $\underline{2}$–cycle. Then $N = \mathcal{S}_n$.

PROOF. Assume N is a normal subgroup of $\mathcal{S}_n$ and it contains a $\underline{2}$–cycle $(i\ j)$. I claim that N contains all $\underline{2}$–cycles. Let $\tau \in \mathcal{S}_n$ be an arbitrary $\underline{2}$–cycle. Since all $\underline{2}$–cycles in $\mathcal{S}_n$ are conjugate $(i\ j)$ and τ are conjugate. Since $(i\ j) \in N$ (and $N \trianglelefteq G$) and $(i\ j)$ and τ are conjugate, $\tau \in N$. As τ was an arbitrary $\underline{2}$–cycle, N must contain all $\underline{2}$–cycles. Since the set of $\underline{2}$–cycles generates $\mathcal{S}_n$ and are contained in N, $N = \mathcal{S}_n$. ($\mathcal{S}_n$ is the smallest group containing all $\underline{2}$–cycles.) ■

9.2.3 EXAMPLES PERTAINING TO $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$

Let p be a positive prime number. Then the determinant 1 matrices

$$\mathfrak{L} \stackrel{\text{def}}{=} \begin{pmatrix} [1]_p & [1]_p \\ [0]_p & [1]_p \end{pmatrix} \quad \text{and} \quad \mathfrak{R} \stackrel{\text{def}}{=} \begin{pmatrix} [1]_p & [0]_p \\ [1]_p & [1]_p \end{pmatrix}$$

generate all of $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$. In symbols:

$$\mathbf{SL}_2\left(\mathbb{Z}_p\right) = \langle\, \mathfrak{L}, \mathfrak{R} \,\rangle\,.$$

(Of course $\mathfrak{L}$ and $\mathfrak{R}$ depend on p and, as such, should really be written as $\mathfrak{L}_p$ and $\mathfrak{R}_p$ since for different primes p the groups $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ have nothing to do with each other.)

CONVENTION 9.15. For ease in this example, I will leave off brackets around all elements of $\mathbb{Z}_p$. All numbers are understood to be the class of the written number.

For starters note that

$$\mathfrak{L}^m = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}$$

and

$$\mathfrak{R}^m = \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix}.$$

Now let

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

be an arbitrary 2 by 2 matrix in $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$. We need to show that it can be written as a product of $\mathfrak{L}$'s and $\mathfrak{R}$'s. This turns out to be a restatement of a fact that most everybody should be familiar with from a linear Algebra class (well ... of those who have taken a linear algebra class ..) that any invertible matrix can be row reduced to a diagonal matrix. We illustrate a technique of row reduction (Guassian elimination).

We define the *row operations* $R_1 \to R_1 + mR_2$ (or written simply as $R_1 + mR_2$ or even as $m\,R_2$ when confined to the 2 by 2 world) as the operation which takes

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longrightarrow \begin{pmatrix} a + mc & b + md \\ c & d \end{pmatrix}$$

and $R_2 \to R_2 + mR_1$ (or $R_2 + mR_1$ or even as $m\,R_2$ when confined to the 2 by 2 world) as the operation which takes

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longrightarrow \begin{pmatrix} a & b \\ c + ma & d + mb \end{pmatrix}.$$

Next we describe an algorithm which reduces any matrix $\mathbf{A} \in \mathbf{SL}_2\left(\mathbb{Z}_p\right)$ to the identity matrix by a sequence of row operations.

ALGORITHM 9.16.

Step 1 Since $\mathbf{A} \in \mathbf{SL}_2\left(\mathbb{Z}_p\right)$ it's first column must be non-zero (any matrix with a zero column must have determinant zero). After possibly performing a $R_1 + R_2$ row operation,[II] we can assume that A_{21} does not equal 0. This output is the output from **Step 1** and also denoted **A**.

Step 2 Since in the output from the previous step we have $A_{21} \neq 0$ we can solve the equation $A_{11} - m\,A_{21} = 1$ for m by taking

$$m = (A_{21})^{-1}(A_{11} - 1).$$

(Every non-zero element of $\mathbb{Z}_p$ has a multiplicative inverse whenever p is prime). Then perform the row operation $R_1 - mR_2$ to the output from the first step. By construction, the new matrix as upper left hand corner equal to 1. The resulting matrix is the output for **Step 2**. As in the previous step, this output is still called **A**.

$$\begin{pmatrix} 1 & * \\ * & * \end{pmatrix}$$

[II] This is short for $R_1 \to R_1 + R_2$

Step 3 Perform the row operation $R_2 - A_{21} R_1$ on the output from **Step 2**. This is the output of **Step 3**. Since the output from the previous step left us with a matrix whose $1,1$ entry was 1, this output has $2,1$ entry o. As before we call this out **A**.

$$\begin{pmatrix} 1 & * \\ 0 & * \end{pmatrix}$$

Step 4 Since every step of this procedure leaves the determinant invariant, the matrix we start this step with must still have determinant 1. Since the first column of the matrix we start this step with is $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$, we must have that the lower right hand entry is also 1, i.e. $A_{22} = 1$. We then perform the row operation $R_1 - A_{12} R_2$ which kills off the entry in the upper right hand side of the matrix. We are then left with the identity matrix.

EXAMPLE 9.17. For instance, let's show how to reduce the matrix

$$\begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}_7).$$ [III]

- Here, since $A_{21} = 2 \neq 0$ **Step 1** is already done for us.
- In $\mathbb{Z}_7$, $2^{-1} = 4$ (since $2 \cdot 4 = 8 \equiv_7 1$)

$$m = (A_{21})^{-1}(1 - A_{11}) = 4 \cdot (1 - 4) = 4 \cdot (-3) \equiv_7 4 \cdot 4$$

  Thus $m = 5$. Then, after doing $R_1 - 5 R_2$ we have

$$\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}_7)$$

- This step leaves us with

$$\mathfrak{L} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}_7).$$

- And, as always, this step leaves us with the identity matrix,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}_7).$$

You should be asking yourself what in the world this has to do with our $\mathfrak{L}$ and $\mathfrak{R}$ matrix friends introduced way at the beginning of this example. The connection is made by noting that for an arbitrary

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}_p)$$

we have that

$$\mathfrak{L}^m \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + mc & b + md \\ c & d \end{pmatrix}$$

and

$$\mathfrak{R}^m \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c + ma & d + mb \end{pmatrix}$$

[III] This really should be written as

$$\begin{pmatrix} [4]_5 & [2]_5 \\ [2]_5 & [3]_5 \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}_7),$$

but if you really think I'm going to write that out 15,000 times you are nuts!

That is we have the following:

$$R_1 \longrightarrow R_1 + mR_2 \rightsquigarrow \text{ left multiplication by } \mathfrak{L}^m$$
$$R_2 \longrightarrow R_2 + mR_1 \rightsquigarrow \text{ left multiplication by } \mathfrak{R}^m.$$

EXERCISE 9.6. What is the relationship between *right* multiplication by $\mathfrak{L}$ and $\mathfrak{R}$, and so-called column operations?

SOLUTION

This finally shows us that $\mathfrak{L}$ and $\mathfrak{R}$ generates $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$. Let's illustrate this with our example in $\mathbf{SL}_2\left(\mathbb{Z}_7\right)$.

$$\mathfrak{L}^{-5} \cdot \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & -5 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$$
$$\Rightarrow \mathfrak{R}^{-2}\,\mathfrak{L}^{-5} \cdot \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$
$$\Rightarrow \mathfrak{L}^{-1}\,\mathfrak{R}^{-2}\,\mathfrak{L}^{-5} \cdot \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Thus, bringing everybody to the other side we have that

$$\begin{aligned}
\begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} &= \left(\mathfrak{L}^{-1}\,\mathfrak{R}^{-2}\,\mathfrak{L}^{-5}\right)^{-1}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
&= \left(\mathfrak{L}^{-1}\,\mathfrak{R}^{-2}\,\mathfrak{L}^{-5}\right)^{-1} && \text{(identity matrix is an identity)} \\
&= \left(\mathfrak{L}^{-5}\right)^{-1}\left(\mathfrak{R}^{-2}\right)^{-1}\left(\mathfrak{L}^{-1}\right)^{-1} && \text{(socks and shoes)} \\
&= \mathfrak{L}^{5}\,\mathfrak{R}^{2}\,\mathfrak{L}^{1} \\
&= \mathfrak{L}^{5}\,\mathfrak{R}^{2}\,\mathfrak{L} \\
&= \begin{pmatrix} 1 & 5 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. && \text{(writing the } \mathfrak{L}\text{'s and } \mathfrak{R}\text{'s out)}
\end{aligned}$$

(1) Another generating set of $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ is

$$\mathbf{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad \Omega = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$$

(where, as usual, we are really considering the classes of the numbers which appear.) For starters, note that both Ω and $\mathbf{J}$ are actually in $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ since they both have determinant 1. Next, we need to show that both $\mathfrak{R}$ and $\mathfrak{L}$ are words in $\mathbf{J}$ and Ω. Indeed, this is the case since a direct calculation shows us that

$$\mathbf{J}^{-1} \cdot \Omega^{-1} = \mathfrak{R}$$

and

$$\mathbf{J} \cdot \Omega = \mathfrak{L}.$$

Thus we have that $\langle\, \mathfrak{L}, \mathfrak{R} \,\rangle = \langle\, \mathbf{J}, \Omega \,\rangle = \mathbf{SL}_2\left(\mathbb{Z}_p\right)$.

(2) It is also true that $\mathfrak{L}$ and $\mathfrak{R}$ generate $\mathbf{SL}_2(\mathbb{Z})$, and we can still follow the same basic outline to prove it. However, **Step 2** needs to be expanded on considerably from its fairly simple form in the case of the $\mathbb{Z}_p$'s. What goes wrong is the fact that we can no longer divide in $\mathbb{Z}$. However, since we know that for a matrix $\mathbf{A} \in \mathbf{SL}_2(\mathbb{Z})$, that A_{11} and A_{21} are relatively prime, (if not their common divisor would divide the determinant which would imply the determinant cannot be one) we can use repeated division algorithm to eventually get $A_{11} = 1$ and $A_{12} = 0$. From there, we can proceed to steps 3 and 4 to finish things off. For instance:

$$\begin{pmatrix} 13 & 5 \\ 5 & 2 \end{pmatrix} \xrightarrow{-2R_2} \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} \xrightarrow{-R_1} \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} \xrightarrow{-R_2} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \xrightarrow{-2R_1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

so

$$\mathfrak{R}^{-2}\,\mathfrak{L}^{-1}\,\mathfrak{R}^{-1}\,\mathfrak{L}^{-2} \cdot \begin{pmatrix} 13 & 5 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 13 & 5 \\ 5 & 2 \end{pmatrix} = \mathfrak{L}^2\,\mathfrak{R}\,\mathfrak{L}\,\mathfrak{R}^2.$$

For another example consider:

$$\begin{pmatrix} 13 & 8 \\ 73 & 45 \end{pmatrix} \xrightarrow{-5R_1} \begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix} \xrightarrow{-R_2} \begin{pmatrix} 5 & 3 \\ 8 & 5 \end{pmatrix}$$
$$\xrightarrow{-R_1} \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} \xrightarrow{-R_2} \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} \xrightarrow{-R_1} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$
$$\xrightarrow{-R_2} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \xrightarrow{-R_1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

which implies that[IV]

$$\begin{pmatrix} 13 & 8 \\ 73 & 45 \end{pmatrix} = \mathfrak{R}^5\,\mathfrak{L}\,\mathfrak{R}\,\mathfrak{L}\,\mathfrak{R}\,\mathfrak{L}\,\mathfrak{R} = \mathfrak{R}^5 \cdot (\mathfrak{L}\,\mathfrak{R})^3.$$

Exercise 9.7. Since $\mathfrak{L}_p$ and $\mathfrak{R}_p$ generate $\mathbf{SL}_2(\mathbb{Z}_p)$, we know that Ω and $\mathbf{J}$ can be written as a product of $\mathfrak{L}$'s and $\mathfrak{R}$'s. Show this directly by explicitly writing i and x as words in $\mathfrak{L}$ and $\mathfrak{R}$.

SOLUTION

Exercise 9.8. Let x be a generator of $\mathcal{U}(\mathbb{Z}_p)$ (see Example 5.75). Show that $\mathbf{GL}_2(\mathbb{Z}_p)$ is generated by $\mathfrak{L}$, $\mathfrak{R}$ and

$$\begin{pmatrix} 1 & 0 \\ 0 & x \end{pmatrix}$$

SOLUTION

[IV] Don't be lazy! Verify these direct matrix computations!

Exercise 9.9. Show that

$$\mathfrak{L} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{F}_{(1\,2)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

generate $^{\pm}\mathbf{SL}_2\left(\mathbb{Z}_p\right)$[V] (for every prime p).

Solution

9.3 Some Exercises

Exercise 9.10. Show that the group of matrices $\mathrm{UT}_3(\mathbb{Z}_n)$ from Exercise 1 on page 334 is generated by

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hint! It may be slightly tricky to see that the matrix

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

can be written as a product of the given two, but after showing this the rest shouldn't be too bad.

Solution

Exercise 9.11. Write the matrices

(1) $\begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}_5)$,

(2) $\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}_7)$,

(3) $\begin{pmatrix} 162 & 31 \\ 47 & 9 \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z})$

in terms of $\mathfrak{L}$ and $\mathfrak{R}$.[I]

Solution

[V] Matrices with determinant either equal to +1 or to −1.

Exercise 9.12. Write the permutation $(1\,5\,3\,7\,6)(2\,4)$ in terms of

Solution

(1) adjacent transpositions

$$(1\,2),(2\,3),(3\,4),(4\,5),(5\,6),(6\,7)$$

(2) the permutations

$$(1\,2) \quad \text{and} \quad (1\,2\,3\,4\,5\,6\,7)$$

(3) the transpositions

$$(1\,2),(1\,3),(1\,4),(1\,5),(1\,6),(1\,7).$$

Exercise 9.13. Let

Solution

$$x := \begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad y := \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}$$

Show that $|x| = |y| = 2$, but that $|xy|$ is infinite. What is the order of the group $\langle x, y \rangle$?
Hint! Use Excercise 3.91.

Exercise 9.14. Suppose that $X, Y \subseteq \mathcal{S}_n$ and that every permutation of X is disjoint from every permutation in Y. Prove that

Solution

$$\langle X \cup Y \rangle \cong \langle X \rangle \times \langle Y \rangle.$$

Example 9.18. We give an example of a use of the previous exercise, Exercise 9.14,

$$\langle (1\,2),(1\,2\,3),(4\,5),(5\,6) \rangle \cong \langle (1\,2),(1\,2\,3) \rangle \times \langle (4\,5),(5\,6) \rangle \cong \mathcal{S}_3 \times \mathcal{S}_3.$$

For ease we now assume that $\langle X \rangle$ is contained in some finite group, G. In what follows it is not too difficult to "fix" in other cases; one typically needs to replace X by $X \cup X^{-1}$ to get the same results proven here.

Definition 9.19. Let W be any subset (not necessarily a subgroup) of a group G. We say that $g \in G$ *preserves* W if $gw \in W$ for all $w \in W$.

Lemma 9.20. *The set of elements of G which preserve a subset $W \subseteq G$ is a subgroup.*

Proof. Since the set of elements which preserve W is a non-empty (it contains the identity) subset of the finite group G we only need to show closure. To that end suppose that g and h preserve W and let $w \in W$ be arbitrary. Since h preserves W, we have that $w' = hw \in W$. But since g also preserves W we must have that gw' is also in W. Hence $ghw \in W$ and gh preserves W. ∎

Remark 9.21. In the case of an infinite group we need to require that $gW = W$ (what we have assumed amounts to simply requiring that $gW \subseteq W$). For instance if we let $\mathbb{N} \subseteq \mathbb{Z}$, the natural numbers, then the set of elements of $\mathbb{Z}$ which preserve this set is the set of natural numbers themselves, which is not a (sub)group.

Theorem 9.22. *If $W \subseteq \langle X \rangle$ is non-empty and preserved by every element of X, then $W = \langle X \rangle$.*

Proof. The assumption tells us that $\langle X \rangle$ preserves W. Let w be an element of W (which exists since $W \neq \emptyset$ by assumption) and let $g \in \langle X \rangle$. Then since $w \in W$ we have that $w \in \langle X \rangle$. Since $\langle X \rangle$ is a group $w^{-1} \in \langle X \rangle$ as well as gw^{-1}. Since $\langle X \rangle$ preserves W we have that $(gw^{-1}) \cdot w = g(w^{-1}w = g$ is still in W. Since g was arbitrary, we have that $\langle X \rangle \subseteq W$. The opposite inclusion is part of the theorems hypothesis. ∎

Corollary 9.23. *Suppose that for some positive integer n we have that*

$$W_{n+1}(X) \subseteq \bigcup_{i=0}^{n} W_i(X)$$

then

$$\langle X \rangle = \bigcup_{i=0}^{n} W_i(X).$$

Proof. Let W be the set

$$\bigcup_{i=0}^{n} W_i(X).$$

By the theorem, we only need to show that W is preserved by every element of X. In general, we have that $xw \in W_{i+1}(X)$ whenever $w \in W_i$ and $x \in X$. Hence for any $w \in W_i(X)$ we have that $xw \in W$ holds for $i < n$ in general. The only remaining case, $w \in W_n(X)$ also works out by the assumption of the corollary. ∎

This corollary gives us an algorithm to calculate $\langle X \rangle$ whenever $\langle X \rangle$ is known to live in some finite group. This is particularly useful in order to calculate $\langle X \rangle$ when X is contained in some symmetric group.

Algorithm 9.24.

Step 1 After deleting elements of X assume that X does not contain the identity or any power of a different element of X. (Hence $\left\langle x, x^2, y, e \right\rangle$ would be replaced by $\langle x, y \rangle$.)

Step 2 Take $W_0 = \{e\}$

Step 3 Assuming we have calculated W_n we define W_{n+1} as the set of

$$\{xw : x \in X, w \in W_n\}$$

Step 4 Delete all elements of W_n which appear in some W_i (for $i < n$).

Step 5 The algorithm ends when some $W_n = \emptyset$.

REMARK 9.25.

- This algorithm always terminates when $\langle X \rangle$ is finite.
- If $\langle X \rangle$ were to end up being infinite this same algorithm could of course be applied, but alas it would never terminate.
- By definition, $W_1 = X$.
- It doesn't matter if the list of words is closed under right multiplication or left multiplication. The same proof applies.

9.4 GENERAL PROPERTIES

PROPOSITION 9.26. *Suppose that*

$$\varphi \colon G \to H$$

is a group homomorphism. Suppose that $X \subseteq G$. Then

(1) $\varphi(\langle X \rangle) = \langle \varphi(X) \rangle$.[I]

(2) *If X generates G, then φ is uniquely determined by what it does to X. More precisely, if φ_2 is also a homomorphism between G and H with $\varphi(x) = \varphi_2(x)$ for every $x \in X$, then $\varphi = \varphi_2$, i.e. if two homomorphisms are the same on some generating set they are the same function.*

REMARK 9.27. To prove this abstractly is a bit annoying, and uses the basic properties of images and inverse images of a set under a function that many undergraduates bumble. Thus we only give the concrete proof with all of its notational nightmares. I recommend that you take a second and work out some specific examples of this to make the notation seem a bit better. The proof of this theorem in many ways hides the fact how obvious this fact really is!

PROOF.

(1) Suppose that $X = \{x_1, x_2, \ldots\}$. Write y_i as the image of x_i under φ. Suppose that $w \in \langle X \rangle$. Then

$$w = x_{a_1}^{n_1} \ldots x_{a_r}^{n_r}$$

for $n_i \in \mathbb{Z}$ and $x_{a_i} \in X$. Then we have

$$\begin{aligned} \varphi(w) &= \varphi\left(x_{a_1}^{n_1} \cdots x_{a_r}^{n_r}\right) \\ &= \varphi\left(x_{a_1}^{n_1}\right) \cdots \varphi\left(x_{a_r}^{n_r}\right) \\ &= \varphi(x_{a_1})^{n_1} \cdots \varphi(x_{a_r})^{n_r} \\ &= y_{a_1}^{n_1} \ldots y_{a_m}^{n_r} \in \langle \varphi(X) \rangle . \end{aligned}$$

Thus, since w was arbitary, we have proven that $\varphi(\langle X \rangle) \subseteq \langle \varphi(X) \rangle$.

[I] In words, the image of a homomorphism is generated by the image of a generating set of the domain.

On the other hand, let $w_H \in \langle \varphi(X) \rangle$ be an arbitrary element of $\langle \varphi(X) \rangle$. Then w_H has the form

$$y_{b_1}^{m_1} \dots y_{b_s}^{m_s}$$

for $y_{b_i} \in \varphi(X)$ and $m_i \in \mathbb{Z}$. But y_{a_i} is simply a notational choice for $\varphi(x_{a_i})$. It's now easy to check that

$$\varphi(x_{b_1}^{m_1} \dots x_{b_s}^{m_s}) = y_{b_1}^{m_1} \dots y_{b_s}^{m_s} = w_H.$$

Thus $w_H \in \varphi(\langle X \rangle)$ proving the other inclusion.

(2) Easy (but annoying) exercise.

■

We can think of the first part of this proposition as saying that the operators φ and $\langle\ \ \rangle$ in some sense commute.

COROLLARY 9.28. *Suppose $\varphi\colon \langle X \rangle \to H$ is a group homomorphism. Then* $\mathrm{Im}(\varphi) = \langle \varphi(X) \rangle$.

COROLLARY 9.29. *Let $\varphi\colon G \to G$ be a homomorphism of groups. If $\varphi(s) = s$ for all s in some generating set of G, then φ must be the identity function, $\mathbf{1}_G$.*

PROOF. The assumption is that φ and the identity function, $\mathbf{1}_G$, do the same thing to every element of some generating set. Therefore, by the second part of proposition, Proposition 9.26**(2)**, the two homomorphisms most be equal. ■

PROPOSITION 9.30. *Suppose we have $\varphi\colon G \to H$ is a group homomorphism. Then* TFAE*:*

(1) *φ is surjective.*

(2) *For every generating set X of G, $\varphi(X)$ generates H.*

(3) *For some generatng set X of G, $\varphi(X)$ generates H.*

(4) *For every generating set Y of H, $Y \subseteq \mathrm{Im}(\varphi)$.*

(5) *For some generating set Y of H, $Y \subseteq \mathrm{Im}(\varphi)$.*

PROOF. By the above proposition we have $\mathrm{Im}(\varphi) = \langle \varphi(X) \rangle$. But φ is surjective iff its image, $\mathrm{Im}(\varphi)$, is H. The equivalence of these three facts is just a rewording of this fact. ■

EXAMPLE 9.31. Consider the homomorphism $\mathcal{S}_4 \to \mathcal{S}_3$ which takes

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

(To actually see there is such a homomorphism is left to you. We will see geometric reasons for this at some point). Then since $\mathcal{S}_3$ is generated by $\{(1\,2), (2\,3)\}$ the proposition implies that this homomorphism is surjective.

LEMMA 9.32. *Suppose H is a finite subgroup of a group G which is generated by some set Y. Then $g \in G$ is in the normalizer of H iff for all $y \in Y$, $gyg^{-1} \in H$.*

PROOF. Since H is finite we have that $g \in \mathrm{N}_G(H)$ iff $gHg^{-1} \subseteq H$. But by assumption, $H = \langle Y \rangle$. Hence by the above,

$$gHg^{-1} = g\langle Y \rangle g^{-1} = \left\langle gYg^{-1} \right\rangle \subseteq \langle Y \rangle = H.$$

■

The following theorem is an incredibly useful theorem in practice. You should think of It as a way which allows us to cut down a **ton** of the work in proving a subgroup of a group is normal.

In the following theorem we X is a set with $\langle X \rangle$ a finite group and $Y \subseteq \langle X \rangle$.

Assuming $Y \subseteq \langle X \rangle$ not $Y \subseteq X$

Theorem 9.33. *Suppose for all $x \in X$ and $y \in Y$ we have that $xyx^{-1} \in \langle Y \rangle$. Then $\langle Y \rangle \trianglelefteq \langle X \rangle$.*

Proof. Let's denote $\langle X \rangle$ as G and $\langle Y \rangle$ as H.

We give an abstract proof (i.e. a proof without using words ...). We want to show that $\mathrm{N}_G(H) = G$. This is because by *Exercise* 8.40 (2) a subgroup is normal in a group iff its normalizer is everything. Furthermore, by Part **(1)** of this same homework question, we know that the normalizer is a subgroup of G. Thus it suffices to show that every element of X is inside of this normalizer (by **gen** 3 this would mean that $G \leqslant \mathrm{N}_G(H)$ as well, so equality must hold). Now we know that for every $x \in X$ that c_x is a homomorphism of G. Hence we have that

$$\begin{aligned} xHx^{-1} &= c_x(\langle Y \rangle) \\ &= \langle c_x(Y) \rangle \\ &= \left\langle xYx^{-1} \right\rangle \\ &\subseteq \langle Y \rangle && \text{(This is the hypothesis of the theorem.)} \\ &= H. && \text{(By definition, } \langle Y \rangle = H\text{)} \end{aligned}$$

That is $xHx^{-1} \subseteq H$, and by **(6)** on page 412 of this same problem on normalizers, $x \in \mathrm{N}_G(H)$ as desired. ∎

Symbolically, the proof of the above is as follows:

$$\begin{aligned} \forall x \in X,\ xYx^{-1} \subseteq \langle Y \rangle &\Rightarrow \forall x \in X,\ x \langle Y \rangle x^{-1} \subseteq \langle Y \rangle \\ &\Rightarrow \forall x \in X,\ x \in \mathrm{N}_G(H) \\ &\Rightarrow X \subseteq \mathrm{N}_G(H) \\ &\Rightarrow \overset{\langle X \rangle}{\overset{\shortparallel}{G}} \leqslant \mathrm{N}_G(H) \\ &\Rightarrow G = \mathrm{N}_G(H) \\ &\Rightarrow H \trianglelefteq G. \end{aligned}$$

Remark 9.34.

- As usual a concrete proof of the above is not difficult, it is just awkward. This is especially the case here. In words, how it would work is that you would write an arbitrary element of $\langle Y \rangle$ in terms of Y and an arbitrary element of the bigger group, $\langle X \rangle$, in terms of X. Then for each conjugation ghg^{-1} pass the first letter of g over the letters of h using the op nature of conjugation. Each time you do this you form a (more than likely) bigger word in $\langle Y \rangle$. What results is the product of elements of $\langle Y \rangle$, and as a group, is an element of $\langle Y \rangle$. We are now left with a completely new element of $\langle Y \rangle$ which we redefine h to be, and a new element of $\langle X \rangle$, which we redefine to be g, which is the same as the original g, but with the last letter hacked off. Now proceed inductively; each step cutting down on the length of g (but probably increasing the length of h, but we don't care). In the end we will have hacked off all letters of g and (as such) are left with identity and an element of H. This should be the intuition behind the Theorem

Theorem 9.33 works because it works one step at a time.

Someday, perhaps a year from now, this will fully click and you will look back and be very annoyed with me! Then, two years from now when you think a bit more of the difference between the proofs you will be probably thank me ⌣̈.

- The only change that would occur if we allowed $\langle X \rangle$ to be infinite is that we would have to also check that $x^{-1}yx \in \langle Y \rangle$ as well.
- Of course there is no need to check pairs of x's and y's that happen to commute because in these cases conjugation doesn't do a damn thing. Also there is no need to check anything about an x which is already in $\langle Y \rangle$ since for such an x we have that $xyx^{-1} \in \langle Y \rangle$ simply by the closure of the group $\langle Y \rangle$. In particular, we won't check the cases when a particular y is a power of x. These cases will be called trivial cases and the rest will be called non-trivial.

Exercise 9.15. In this exercise, I have you give two proofs of the following: For any subset $X \subseteq G$

$$C_G(X) = C_G(\langle X \rangle).^{\text{II}}$$

SOLUTION

In particular, since $C_G(G) = \mathcal{C}(G)$, the center of G, this says that in words, an element $g \in G$ commutes with every element of G iff it commutes with every element in some generating set of G.

(1) To give the concrete proof follow the following outline: assume that $g \in G$ commutes with every $x \in X$, then

- Show that assumption guarantees that g commutes with the inverse of every element of X. (What you're really showing is that in general g commutes with $h \in G$ implies that g commutes with h^{-1}. We've done this before.)
- Show that the assumption guarantees that g commutes with any power of any element of X (the 1^{st} powers are by assumption, the -1 power was the above bullet point).
- Let $w \in G$ be an arbitrary element (hence a word in X). Use the above bullet to show that g commutes with w by writing w out as

$$x_1^{n_1} \cdots x_m^{n_m}$$

and "dragging" (commuting) g by each $x_i^{n_i}$ individually . Perhaps an induction will make this "look" a little cleaner.

(2) For the abstract proof (again assuming that some $g \in G$ commutes with every element of X.

- Show that the assumption that every element of X commutes with g implies that $X \subseteq C_G(g)$ (again, this completely obvious).
- Conclude from the definition of $\langle X \rangle$ and the original assumption that $\langle X \rangle = G$ to prove that $G \leqslant C_G(g)$.
- Conclude that $G = C_G(g)$.
- Use Exercise 6.11 back on page 326 to conclude that $g \in \mathcal{C}(G)$.

Example 9.35.

(1) Let $G = \mathcal{D}_4 (= \langle \rho, \mathbf{r} \rangle)$ and $H = \langle \rho \rangle = \{\mathbf{1}, \rho, \rho^2, \rho^3\}$. The only non-trivial case to check is $\mathbf{r}\rho\mathbf{r}^{-1}$. But this equals ρ^{-1} which (as an inverse of the element ρ) is an element of $\langle \rho \rangle$. Hence H is a normal subgroup of $\mathcal{D}_4$. If we didn't have this theorem, and we did this by definition we would have to check all 8 elements of $\mathcal{D}_4$ off of all 4

[II] This says, among other things that G is abelian iff any two elements of some generating set commute.

elements of H for a total of 32 cases (or 21 cases if we ignore the cases involving the identity) versus the 1 case that we checked using the theorem.

(2) We consider the binary tetrahedra group, $B(\blacktriangleright^3)$, first considered in {where }. In a few pages, in Section 9.7, we will show that $B(\blacktriangleright^3)$ is generated as follows:

$$B(\blacktriangleright^3) = \left\langle \mathbf{i}, \overset{\omega}{\overbrace{\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}}} \right\rangle \leqslant \mathbb{S}^3.$$

In a long winded example, Example 3.113, we showed that $\omega\mathbf{i}\omega^{-1} = \mathbf{j}$ and $\omega\mathbf{j}\omega^{-1} = \mathbf{k}$. In particular we have that $\mathbf{j} \in B(\blacktriangleright^3)$ and so $Q_8 = \langle \mathbf{i}, \mathbf{j} \rangle \leqslant G$. Although $\langle \mathbf{i} \rangle$ isn't normal in $B(\blacktriangleright^3)$ (it is normal in Q_8 however), we claim that the larger group, Q_8, is indeed normal in $B(\blacktriangleright^3)$. Indeed, the only non-trivial cases (any case involving $\mathbf{i}$ is trivial since $\mathbf{i} \in Q_8$) that we need to check are $\omega\mathbf{i}\omega^{-1}$ and $\omega\mathbf{j}\omega^{-1}$. But the first is equal to $\mathbf{j}$ and the second is equal to $\mathbf{k}$, both of which are in Q_8. Hence Q_8 is normal in $B(\blacktriangleright^3)$, i.e. $Q_8 \trianglelefteq B(\blacktriangleright^3)$.

(3) We are about to see that

$$B(\blacklozenge^3) = \left\langle \overset{S_\mathbf{i}}{\overbrace{\frac{1}{\sqrt{2}}(1+\mathbf{i})}}, \overset{S_\mathbf{j}}{\overbrace{\frac{1}{\sqrt{2}}(1+\mathbf{j})}} \right\rangle.$$

Then we have $S_\mathbf{i} \cdot S_\mathbf{j} = \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}$ and $S_\mathbf{i}^2 = \mathbf{i}$. Since these two quaternions generate $B(\blacktriangleright^3)$ (see last part) we have $B(\blacktriangleright^3) \leqslant B(\blacklozenge^3)$. In particular $Q_8 \leqslant B(\blacklozenge^3)$. I claim that once again Q_8 is normal, this time in $B(\blacklozenge^3)$. Indeed, we have

c_g	$S_\mathbf{i}$	$S_\mathbf{j}$
$\mathbf{i}$	$S_\mathbf{i}\,\mathbf{i}\,S_\mathbf{i}^{-1} = \mathbf{i}$	$S_\mathbf{j}\,\mathbf{i}\,S_\mathbf{j}^{-1} = -\mathbf{k}$
$\mathbf{j}$	$S_\mathbf{i}\,\mathbf{j}\,S_\mathbf{i}^{-1} = \mathbf{k}$	$S_\mathbf{j}\,\mathbf{j}\,S_\mathbf{j}^{-1} = \mathbf{j}$

Since every entry of the table ends up being an element of Q_8 we have proved that $Q_8 \trianglelefteq B(\blacklozenge^3)$.

(4) We are about to see that (actually, we are going to reference a theorem from the appendix to do this)

$$B(\bigstar^3) = \left\langle \mathbf{i}, \overset{q_5}{\overbrace{\frac{1}{2}\phi + \frac{1}{2}\mathbf{i} + \frac{1}{2}\phi^{-1}\mathbf{j}}} \right\rangle.$$

With a bit of effort, we can see that $\mathbf{j} \in B(\bigstar^3)$, in particular, since both $\mathbf{i}$ and $\mathbf{j}$ are in $B(\bigstar^3)$, we again have $Q_8 \leqslant B(\bigstar^3)$. However this time, unlike the two previous times we do not have $Q_8 \trianglelefteq B(\bigstar^3)$. Indeed

$$q_5\,\mathbf{i}\,q_5^{-1} = \frac{1}{2}\phi\,\mathbf{i} + \frac{1}{2}\phi^{-1}\,\mathbf{j} + \frac{1}{2}\mathbf{k}$$

which is not an element of $Q_8 = \{\pm 1, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}\}$.

(5) Consider $G = \mathcal{S}_4$ and V the Klein 4-group,

$$\left\langle \overset{a}{\overbrace{(1\,2)(3\,4)}}, \overset{b}{\overbrace{(1\,4)(2\,3)}} \right\rangle.$$

The conclusion of the theorem does not require we use a particular generating set, so we will get the same answer no matter the generating set we choose for $\mathcal{S}_4$. My favorite generating set of $\mathcal{S}_4$ is $\{s_1 = (1\,2), s_2 = (2\,3), s_3 = (3\,4)\}$. We have that s_1 and s_3 commute with a while y commutes with b. Thus the non-trivial cases to check are xbx^{-1}, zbz^{-1}, and yay^{-1} but all 3 of these cases equal $(1\,3)(2\,4)$ which happens to equal a times b, ab, which, lucky for us , is an element of V by closure.

(6) Let

$$x = (2\ 10\ 6)(4\ 8\ 12)$$
$$y = (1\ 10)(2\ 5)(4\ 7)(6\ 9)(8\ 11)$$
$$z = (1\ 7)(2\ 8)(3\ 9)(4\ 10)(5\ 11)(6\ 12)$$

and consider the group $\langle x, y, z \rangle$, which we denote as G. For starters, note that y and z commute. As usual in $\mathcal{S}_n$, the easiest way to check this (since conjugating is so much easier than composing in permutation groups) is to note that $yzy^{-1} = z$. (To help your intuition, it could be helpful to draw the cycle diagrams for both y and z at the same time and notice that this diagram is cut into little squares). We also have

$$yxy^{-1} = (5\ 1\ 9)(7\ 11\ 3)(1\ 9\ 5)(3\ 7\ 11)$$

which we call w and

$$zxz^{-1} = (8\ 4\ 12)(10\ 11\ 3) = x^{-1}$$
$$zwz^{-1} = zyxy^{-1}z^{-1} = yzxz^{-1}y^{-1} = yx^{-1}y^{-1} = w^{-1}.$$

Notice that w is a $\underline{3}\,\underline{3}$-cycle disjoint from (hence commuting with) x. Now we consider $N = \langle x, w \rangle$. The table coupled with the theorem shows that $N \trianglelefteq G$

c_g	x	y	z
x	$xxx^{-1} = x$	$yxy^{-1} = w$	$zxz^{-1} = x^{-1}$
w	$xwx^{-1} = w$	$ywy^{-1} = x$	$zwz^{-1} = w^{-1}$

.

Exercise 9.16. Consider a group G which is generated by two elements x and y. Suppose further that $yxy^{-1} = x^i$ for some integer i. Show that both $\langle X \rangle \trianglelefteq G$ and $\left\langle x^2 \right\rangle \trianglelefteq G$. Find an example of this where $|G| = |x| \cdot |y|$ and a (necessarily different) example of this where $|G| \neq |x| \cdot |y|$.
Hint! One can look just at groups of order 8 for both kinds of examples.

SOLUTION

Exercise 9.17. Now we let G be the group

$$\left\langle (1\ 2\ 3\ \overset{x}{\overset{\shortparallel}{4}})(6\ 7), (1\ 3)(2\ \overset{y}{\overset{\shortparallel}{4}})(5\ 6\ 7) \right\rangle$$

Show that $\langle x \rangle$ is not normal in G, but $\langle x^2 \rangle$ and $\langle y \rangle$ are normal in G.

SOLUTION

EXERCISE 9.18. If $G = \langle x, y \rangle$ prove that

$$\left\langle x^2, xy, y^2 \right\rangle \trianglelefteq G.$$

SOLUTION

Hint! Start off by proving that yx is in this subgroup.

EXAMPLE /THEOREM 9.36. Suppose that

$$G = \langle x_1, \ldots, x_m \rangle$$

is an abelian group. Then the homomorphism from Theorem 8.75

$$\begin{aligned} &: \mathbb{Z}_{|x_1|} \times \cdots \times \mathbb{Z}_{|x_m|} \to G \\ &: (a_1, \ldots, a_m) \mapsto {x_1}^{a_1} \cdots {x_m}^{a_m} \end{aligned}$$

is surjective and is an isomorphism iff

$$|G| = |x_1| \cdots |x_m|.$$

PROOF. For notational ease, let's denote $|x_i|$ as n_i. Similar to Theorem 8.75 we have a homomorphism of groups

$$\begin{aligned} \varphi &: \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_m} \to G \\ &: ([a_1]_{n_1}, \ldots, [a_m]_{n_1}) \mapsto x_1^{a_1} \cdots x_m^{a_m}. \end{aligned}$$

As pointed out after the proof of that theorem, the image of this homomorphism is

$$\langle x_1 \rangle \langle x_2 \rangle \cdots \langle x_m \rangle.$$

But since each $\langle x_i \rangle \trianglelefteq G$, we know that this product is $\langle x_1, \ldots, x_m \rangle$, which, by assumption, is all of G. Our other assumption is that

$$|G| = n_1 \cdots n_m.$$

But this is precisely the order of the domain of φ. As such φ is a surjective homomorphism between two groups of the same size, which, by the pigeon hole principle, must be an isomorphism. ■

Of course this lemma is not true if A didn't happen to be abelian. For instance $\mathcal{S}_3$ is generated by elements of order 2 but contains an element of order 3.
The following notation is so obvious it almost doesn't need to be given:

NOTATION 9.37. Let $\mathcal{A}$ be a subset of a group. Define $\operatorname{lcm}(|\mathcal{A}|)$ as

$$\operatorname{lcm}(\{|a| : a \in \mathcal{A}\}).$$

Exercise 9.19. Suppose an abelian group A is generated by some set $\mathcal{A}$, i.e. suppose

$$A = \langle \mathcal{A} \rangle .$$

Prove that

$$A = A[\text{lcm}(|\mathcal{A}|)].$$

Hint! Once you recall what the definition of $A[m]$ is, this exercise should take only one line.

SOLUTION

Here's a corollary to this exercise:

Corollary 9.38. *If A is a finite abelian group generated by elements of order p (where p is prime), then every non-identity element of order A has order p.*

Note that this previous corollary is (quite) false if A weren't assumed abelian. For instance the group $\mathcal{S}_3$ (more generally $\mathcal{S}_n$) is generated by elements of order 2, namely $(1\ 2)$ and $(2\ 3)$, but it contains an element of order 3, namely $(1\ 2\ 3)$.

9.5 Normal Closure*

We now define the analogous notion of the *normal closure* of a set, denoted by $\lfloor X \rceil$. Roughly, it is the same definition as the definition of $\langle X \rangle$. Simply replace all occurences of $\leqslant$ with $\trianglelefteq$ in the definition. It can be thought of as the *smallest normal subgroup of G which contains X*.

Definition 9.39. Suppose $X \subseteq G$. $\lfloor X \rceil$ is the unique subgroup of G which satisfies the following three properties:

ngen 1 $X \subseteq \lfloor X \rceil$

ngen 2 $\lfloor X \rceil \trianglelefteq G$.

ngen 3 If $N \trianglelefteq G$, then $X \subseteq N \Rightarrow \lfloor X \rceil \leqslant N$.

Remark 9.40. If X happens to be a subgroup, this is, in a sense, the opposite of the core of X in the sense that the core is the largest normal subgroup of G contained inside of X while the normal closure is the smallest normal subgroup of G containing G. In this way we could have called the core the *normal interior*.

The following exercise shows that $\lfloor X \rceil$ is a well defined concept.

Exercise 9.20.

(1) Prove that the intersection of all normal subgroups of G which contains X satisfies **ngen 1**, **ngen 2**, and **ngen 3**.

SOLUTION

(2) Prove that $\lfloor X \rceil$ is unique. That is prove that two subgroups which satisfy **ngen 1**, **ngen 2**, and **ngen 3** are equal.
Hint! Logically, you should be getting tired of doing proofs like this!

EXERCISE 9.21. Prove that $\langle X \rangle \subseteq \lfloor X \rceil$.

SOLUTION

EXERCISE 9.22. Show by example, unlike $\langle X \rangle$, $\lfloor X \rceil$ depends on G (and so it really should be denoted $\lfloor X \rceil_G$).
Hint! It's hard to mess this up. Pick your favorite group and your favorite non-normal subgroup of that group.

SOLUTION

EXERCISE 9.23. Define $WW(X)$ inductively as follows:

- $e_G \in WW(X)$
- If $w \in WW(X)$, $x \in X$, and $g \in G$, then $wgxg^{-1}$ and $wgx^{-1}g^{-1}$ are also in $WW(X)$.

Prove that $\lfloor X \rceil = WW(X)$.

SOLUTION

EXERCISE 9.24. Prove that

$$\lfloor g \rceil = \langle \text{cc}(g) \rangle .$$

SOLUTION

This last exercise implies that

$$\lfloor (1\ 2) \rceil = \langle \text{cc}_{\mathcal{S}_n}((1\ 2)) \rangle = \langle \text{All 2-cycles} \rangle = \mathcal{S}_n.$$

EXERCISE 9.25. Prove for any homomorphism, φ, of G onto H we have $\varphi(\lfloor X \rceil) = \lfloor \varphi(X) \rceil$.

SOLUTION

Rather surprisingly, we will see that $\lfloor X \rceil$ is often easier to calculate then its cousin, $\langle X \rangle$.

9.6 AUTOMORPHISMS

In Exercise 8.10 we defined an automorphism of a group G; it's simply an isomorphism of G to itself. There we showed the set of these guys, denoted as $\text{Aut}(G)$, forms a group.

Definition 9.41. A subgroup $N \leqslant G$ is *characteristic* in G if

$$\forall \varphi \in \mathrm{Aut}(G),\ \varphi(N) = N.$$

If N is characteristic in G, we write N **char** G.

Remark 9.42.

(1) We could have weakened this definition to require that for all automorphisms φ that $\varphi(N) \subseteq N$ or we could have required that $N \subseteq \varphi(N)$. This is because if $\psi(N) \subseteq N$ for all automorphisms ψ, we would have that $\varphi^{-1}(N) \subseteq N$ for all automorphisms ψ (since the set of automorphisms is a group and for any group the set of inverse elements is the same as the set of elements). But then for all automorphisms ψ we would have that

$$N = \varphi(\varphi^{-1}(N)) \subseteq \varphi(N)$$

implying that $N = \varphi(N)$ for all automorphisms N. This is just basic set theory, see Exercise 2.31 back on page 35.

(2) Note that if N **char** G, then for any automorphism φ of G, the restricted map $\varphi|_N$ is an automorphism of N. Indeed, the facts that $\varphi|_N$ is still o.p. and injective are true of any restricted function, while the fact that this restriction is also surjective is just another way of stating that $\varphi(N) = N$. In fact we have that

$$\begin{aligned} &: \mathrm{Aut}(G) \to \mathrm{Aut}(N) \\ &: \varphi \mapsto \varphi|_N \end{aligned}$$

is a homomorphism of groups. (However, this homomorphism needn't be injective or surjective.)

Proposition 9.43. N **char** $G \Rightarrow N \trianglelefteq G$.

Proof. The result follows at once when we realize that for any $g \in G$, conjugation by g, c_g is an automorphism. We leave the details to you in Exercise 9.26. ∎

Exercise 9.26. Fill in the details of this statement, i.e. prove

$$N \textbf{ char } G \Rightarrow N \trianglelefteq G.$$

SOLUTION

The converse, however, is false, i.e. there are normal subgroups of a group which are not characteristic. For instance in $\mathbb{Z}_2^2$, we have that $\langle ([1]_2, [0]_2) \rangle$ is a normal subgroup (it's a subgroup of an abelian group) but it's not characteristic. This is because the map

$$([a]_2, [b]_2) \mapsto ([b]_2, [a]_2)$$

is an automorphism of $\mathbb{Z}_2^2$, which takes $\langle ([1]_2, [0]_2) \rangle$ to $\langle ([0]_2, [1]_2) \rangle$, not itself.

One cool thing about characteristic subgroups is that the relation of being a characteristic subgroup is transitive. Compare this with Exercise 8.42 back on page 412 which showed that the relation $\trianglelefteq$ was **not** transitive. Actually, this is even stronger than mere transitivity.

EXERCISE 9.27. Suppose

$$N_1 \textbf{ char } N_2 \textbf{ char } G$$

prove that N_1 **char** G.

SOLUTION

This exercise is along a similar line as the previous.

EXERCISE 9.28. Suppose

$$N_1 \textbf{ char } N_2 \trianglelefteq G$$

prove that $N_1 \trianglelefteq G$.

SOLUTION

Recall for every $g \in G$, the map

$$h \mapsto ghg^{-1}$$

is an automorphism, which we decided to denote as c_g. In Example/Theorem 8.29 we showed that the assignment

$$g \mapsto c_g$$

gave us a homomorphism $G \to \text{Aut}(G)$ with kernel equal to the center of G. Thus, this homomorphism is injective iff G is centerless. The image of this homomorphism, i.e. the set of all c_g as g varies through the group G, is denoted $\text{Inn}(G)$. It consists of *inner automorphisms*.

Therefore an automorphism $\varphi\colon G \to G$ is an inner automorphism iff there exists $g \in G$ with $\varphi = c_g$. The easiest way of showing that a given automorphism φ is not inner (such automorphism are typically called *outer*) is by pointing to some $k \in G$ with k and $\varphi(k)$ happen to be two non- conjugate elements of G.[I]

One way to think of N **char** G is to think that N is invariant under all automorphisms of G. In this vein, we think of $N \trianglelefteq G$ as the statement that N is invariant under (just) all inner-automorphisms of G. In this way, a fancy way of rationalizing the fact that $\trianglelefteq$ is as follows:

> Inner automorphisms of G do not restrict to *inner* automorphisms of N, even when N is characteristic in G (they do restrict to automorphisms, just not necessarily inner).

For instance, the inner automorphism of $\mathcal{D}_4$ given by conjugation by the reflection $\mathbf{r}$, induces the map

$$: \langle \rho \rangle \to \langle \rho \rangle : \qquad \rho^i \mapsto \rho^{-i}$$

which is **not** an inner automorphism of $\langle \rho \rangle$. (Indeed, since $\langle \rho \rangle$ is abelian, the only inner automorphism of this group is the trivial map.)

Despite how much nicer the properties are for **char** are compared to $\trianglelefteq$ it is often quite difficult to find automorphisms which are not inner. Moreover, when studying quotient groups (starting in Chapter 11 coming up on page 597) we will only need to require that $N \trianglelefteq G$. It is for these reasons (among others) that normal subgroups play a more central role in group theory.

Despite the fact that it can be a rather large challenge to find non-inner automorphisms of a group, here's a fairly simple way of proving that a given subgroup of a group is characteristic.

PROPOSITION 9.44. *Suppose $N \leqslant G$ has the property that N is not isomorphic to any subgroup of G except itself. Then N* **char** *G.*

[I] There are examples where even this doesn't work! That is there are automorphisms φ which take all elements of G to a conjugate element which never the less is not an inner automorphism.

Remark 9.45. Since if $N \cong \widehat{N}$, $|N| = |\widehat{N}|$ we only need to check that N is not isomorphic to any subgroup of G whose order is the same as N's.

Proof. For an arbitrary automorphism φ of G, the groups N and $\varphi(N)$ are isomorphic subgroups of G. By assumption, this means that $\varphi(N) = N$. Since φ was arbitrary, N **char** G. ∎

Example 9.46.

(1) For a finite cyclic group C, there exists a unique subgroup of C whose order is d (for every divisor of C). In particular, every subgroup of a finite cyclic group is characteristic. This is also true for the infinite cyclic group $\mathbb{Z}$, but the reasoning is more subtle.

Note that any two non-trivial subgroups of are isomorphic to each other (all are of the form $n\mathbb{Z}$ and $n\mathbb{Z} \cong m\mathbb{Z}$), but despite all this, there is no automorphism of $\mathbb{Z}$ which gives rise to these isomorphisms! Indeed, the only automorphisms of the integers are the identity (multiplication by 1) and multiplication by -1; both of which take every subgroup of $\mathbb{Z}$ to itself.

The converse of Proposition 9.44 is false.

(2) For example, the group $\langle \rho \rangle \preccurlyeq \mathcal{D}_4$ is the only subgroup of $\mathcal{D}_4$ which is isomorphic to $\mathbb{Z}_4$. (Indeed, it contains the only elements of order 4 in $\mathcal{D}_4$, namley ρ and ρ^{-1}.) Therefore we have that

$$\langle \rho \rangle \textbf{ char } \mathcal{D}_4.$$

(3) For any group G, the center of G,

$$\mathcal{C}(G) \textbf{ char } G.$$

This is because for all $\varphi \in \mathrm{Aut}(G)$, $g \in G$, and $z \in G$ there exists a $\widetilde{g} \in G$ with $\varphi(\widetilde{g}) = g$. This is because automorphisms are surjective. Therefore we have

$$\begin{aligned} \varphi(z)g &= \varphi(z)\varphi(\widetilde{g}) \\ &= \varphi(z\widetilde{g}) && \text{(Since } \varphi \text{ is o.p.)} \\ &= \varphi(\widetilde{g}z) && \text{(Since } z \in \mathcal{C}(G)\text{, } z \text{ commutes with } \widetilde{g}) \\ &= \varphi(\widetilde{g})\varphi(z) \\ &= g\varphi(z). \end{aligned}$$

Therefore $\varphi(z)$ commutes with every element of G and is therefore (still) in the center of G.

(4) The automorphism from Exercise 8.3 which was given by

$$\varphi := \begin{pmatrix} \mathbf{1} & \rho & \rho^2 & \rho^3 & \mathbf{r} & \mathbf{r}\rho & \mathbf{r}\rho^2 & \mathbf{r}\rho^3 \\ \mathbf{1} & \rho & \rho^2 & \rho^3 & \mathbf{r}\rho & \mathbf{r}\rho^2 & \mathbf{r}\rho^3 & \mathbf{r} \end{pmatrix}$$

is not inner. This is because $\varphi(\mathbf{r}) = \mathbf{r}\rho$ but $\mathbf{r}$ and $\mathbf{r}\rho$ are two non-conjugate elements of $\mathcal{D}_4$. Note also this map φ takes the subgroup

$$\varphi\colon \{\mathbf{1}, \mathbf{r}, \mathbf{r}\rho^2, \rho^2\} \to \{\mathbf{1}, \mathbf{r}\rho, \mathbf{r}\rho^3, \rho^2\}$$

which proves that, despite both are $\trianglelefteq \mathcal{D}_4$, neither **char** $\mathcal{D}_4$.

However, note that on the other hand $\varphi^2 = \varphi \circ \varphi$, which is

$$\varphi^2 = \begin{pmatrix} \mathbf{1} & \rho & \rho^2 & \rho^3 & \mathbf{r} & \mathbf{r}\rho & \mathbf{r}\rho^2 & \mathbf{r}\rho^3 \\ \mathbf{1} & \rho & \rho^2 & \rho^3 & \mathbf{r}\rho^2 & \mathbf{r}\rho^3 & \mathbf{r} & \mathbf{r}\rho \end{pmatrix},$$

is an inner automorphism. Indeed it is the same as conjugation by ρ, i.e. $c_\rho = \varphi^2$.

We don't really use this, but understanding the proof could help you understand the concept of the automorphism group better as well give you more practice proving that a subgroup of a group is normal.

Example /Theorem 9.47. $\mathrm{Inn}(G) \trianglelefteq \mathrm{Aut}(G)$.

Proof. Let $\varphi : G \to G$ be an automorphism of G and let $c_g : G \to G$ be an inner automorphism. Then, I claim that

$$\varphi \circ c_g \circ \varphi^{-1} = c_{\varphi(g)}.$$[II]

As usual, when we are attempting to prove the equality of two functions, we simply show that they have the same effect on every point in their domain, which in our current situation, is the group G. Let $a \in G$ be arbitrary. Then

$$\begin{aligned} (\varphi \circ c_g \circ \varphi^{-1})(a) &= \varphi(c_g(\varphi^{-1}(a))) \\ &= \varphi(g \cdot \varphi^{-1}(a) \cdot g^{-1}) \\ &= \varphi(g) \cdot \varphi(\varphi^{-1}(a)) \cdot \varphi(g^{-1}) \\ &= \varphi(g) \cdot a \cdot \varphi(g)^{-1} \\ &= c_{\varphi(g)}(a). \end{aligned}$$

This proves that $\varphi \circ c_g \circ \varphi^{-1} = c_{\varphi(g)}$ and completes the proof that $\mathrm{Inn}(G)$ is normal in $\mathrm{Aut}(G)$ since we've shown that if we conjugate an arbitrary inner automorphism we still have an inner automorphism. ∎

9.6.1 Automorphisms of $\mathbb{Z}_n$

Theorem 9.48. $\mathrm{Aut}(\mathbb{Z}_n^m) \cong \mathbf{GL}_m(\mathbb{Z}_n)$.

Proof. Notice that an automorphism of $\mathbb{Z}_n^m$ is a bijective linear map,

$$\mathbb{Z}_n^m \to \mathbb{Z}_n^m.$$

Conversely, since the operation is addition, such a bijective linear map corresponds to an automorphism of $\mathbb{Z}_n^m$. ∎

Exercise 9.29. Suppose that $\varphi : \widetilde{G} \to G$ is an isomorphism. Prove that for all $g \in \widetilde{G}$,

$$\varphi(cc_{\widetilde{G}}(g)) = cc_G(\varphi(g)).$$

SOLUTION

Hint! Really, what you need to show is that g is conjugate to g' iff $\varphi(g)$ is conjugate to $\varphi(g')$. (**Why?**)

[II] Yes, we are conjugating conjugation! Use your LEFT brain! LEFT brain, this is EASY, don't panic!

COROLLARY 9.49. $\mathrm{Aut}(\mathbb{Z}_n) \cong \mathcal{U}(\mathbb{Z}_n)$.

PROOF. This follows directly from the theorem since $\mathbf{GL}_1(\mathbb{Z}_n) = \mathcal{U}(\mathbb{Z}_n)$. But to be a good guy, I give a direct proof of this. Define a map

$$\begin{aligned}\Phi\colon \mathrm{Aut}(\mathbb{Z}_n) &\to \mathcal{U}(\mathbb{Z}_n)\\ : \varphi &\mapsto \varphi([\,1\,]_n).\end{aligned}$$

I claim that Φ is an isomorphism.

(1) First I need to show that the image of Φ is really contained in $\mathcal{U}(\mathbb{Z}_n)$, as indicated. That is for every automorphism of $\mathbb{Z}_n$, I need to show that $\varphi([\,1\,]_n) \in \mathcal{U}(\mathbb{Z}_n)$. Indeed, this is the case because the (additive) order an element of group after any isomorphism has been applied to it (hence an automorphism) is the same both before and after the application. Thus, since $\varphi\mathrm{Aut}(\mathbb{Z}_n)$, $\varphi([\,1\,]_n)$ must have the same order as $[\,1\,]_n$ does. But $[\,1\,]_n$ has (additive) order n in $\mathbb{Z}_n$, hence so does $\varphi([\,1\,]_n)$. This means $\langle\, \varphi([\,1\,]_n)\,\rangle = \mathbb{Z}_n$ (here I mean things in the additive sense). This means in particular, $[\,1\,]_n \in \langle\, \varphi([\,1\,]_n)\,\rangle$, i.e. there exists an integer m so that $m\varphi([\,1\,]_n) = [\,1\,]_n$. Well this means that $\varphi([\,1\,]_n)$ is invertible (with inverse m) and so $\varphi([\,1\,]_n) \in \mathcal{U}(\mathbb{Z}_n)$.

(2) Next I must show that Φ is operation preserving. ⟨Note that the operation in the first group is composition of functions which preserve addition, and the operation in the second group is multiplication.⟩ Indeed, let φ and ψ be arbitrary automorphisms of $\mathbb{Z}_n$ (i.e. elements of the domain of Φ). Then since φ is an additive automorphism, for any $m \in \mathbb{Z}$ we have that

$$\begin{aligned}\varphi([\,m\,]_n \cdot [\,a\,]_n) &= \varphi(m\,[\,a\,]_n)\\ &= m\varphi([\,a\,]_n)\\ &= [\,m\,]_n\,\varphi([\,a\,]_n).\end{aligned} \tag{1}$$

As such,

$$\begin{aligned}\Phi(\varphi\circ\psi) &= (\varphi\circ\psi)([\,1\,]_n)\\ &= \varphi(\psi([\,1\,]_n))\\ &= \varphi(\psi([\,1\,]_n \cdot [\,1\,]_n))\\ &= \psi([\,1\,]_n)\cdot\varphi([\,1\,]_n) && ((1)\text{ w/ }[\,m\,]_n = \psi([\,1\,]_n), a = 1)\\ &= \varphi([\,1\,]_n)\cdot\psi([\,1\,]_n) && (\text{Mult in } \mathbb{Z}_n \text{ is commutative})\\ &= \Phi(\varphi)\cdot\Phi(\psi).\end{aligned}$$

Therefore Φ is operation preserving.

(3) Lastly we prove that Φ is a bijection by finding an explicit inverse. Namely for any $a \in \mathcal{U}(\mathbb{Z}_n)$ define the map m_a as multiplication by a. This a homomorphism by the distributive property and it has an inverse given by $m_{a^{-1}}$. Hence it is an isomorphism from $\mathbb{Z}_n$ to itself, i.e. an automorphism of $\mathbb{Z}_n$.

Now I claim that the function

$$\begin{aligned}\Psi\colon \mathbb{Z}_n &\to \mathrm{Aut}(\mathbb{Z}_n)\\ a &\mapsto m_a\end{aligned}$$

is an inverse of Φ. Indeed, we have

$$\begin{aligned}\Phi\circ\Psi(a) &= \Phi(\Psi(a))\\ &= \Phi(m_a)\\ &= m_a(1)\\ &= a\cdot 1\\ &= a.\end{aligned}$$

Hence $\Phi \circ \Psi = \mathbf{1}_{\mathcal{U}(\mathbb{Z}_n)}$. Now we still need to show that $\Psi \circ \Phi$ is the identity. Indeed, for any automorphism φ of $\mathbb{Z}_n$, we have that $\Psi \circ \Phi(\varphi) = m_{\varphi([\,1\,]_n)}$. But for every $[\,a\,]_n \in \mathbb{Z}_n$ ($a \in \mathbb{Z}$) we have that

$$\begin{aligned}\varphi([\,a\,]_n) &= \varphi(a \cdot [\,1\,]_n)\\ &= a\varphi([\,1\,]_n)\\ &= [\,a\,]_n \cdot \varphi([\,1\,]_n)\\ &= \varphi([\,1\,]_n) \cdot [\,a\,]_n\\ &= m_{\varphi([\,1\,]_n)}([\,a\,]_n).\end{aligned}$$

Hence we do indeed have the equality $m_{\varphi([\,1\,]_n)} = \varphi$. Therefore $\Psi \circ \Phi$ is the identity and we have established that Ψ and Φ are inverses to each other. ■

We have really shown that every automorphism of $\mathbb{Z}_n$ is multiplication by a for some invertible $a \in \mathbb{Z}_n$.

9.6.2 Automorphisms of $\mathcal{S}_n$

Theorem 9.50. *Suppose that $\varphi\colon \mathcal{S}_n \to \mathcal{S}_n$ is an automorphism which takes 2–cycles to 2–cycles. Then φ is an inner automorphism.*

Proof. Let φ be as in the statement of the theorem. By Exercise 9.3 back on page 442, the set of 2–cycles

$$\{(1\,2), (1\,3), \ldots (1\,n)\} = \{(1\,i) : i = 2, \ldots n\}$$

generates $\mathcal{S}_n$. Therefore, by Proposition 9.26**(2)**, to prove the theorem we need to find an $\sigma \in \mathcal{S}_n$ with

$$\forall i = 2, \ldots, n,\ c_\sigma((2\,i)) = \varphi((2\,i)). \tag{2}$$

By assumption, φ takes 2–cycles to 2–cycles, in particular φ must take every element of X to a 2–cycle. That is, for each $i \in \{2, 3, \ldots n\}$, $\varphi((1\,i))$ is a 2–cycle. So there are integers $a_2, \ldots, a_n$ and $b_2, \ldots, b_n$ (with $a_i \neq b_i$) with

$$\varphi((1\,2)) = (a_2\ b_2), \quad \varphi((1\,3)) = (a_3\ b_3) \ldots \varphi((1\,n)) = (a_n\ b_n)$$

i.e.

$$\varphi((1\,i)) = (a_i\ b_i).$$

Now since $(1\,2)$ and $(1\,3)$ do not commute, we must have that their images under the automorphism φ also do not commute. Hence $\{a_2, b_2\}$ and $\{a_3, b_3\}$ cannot be disjoint. WLOG we assume that $a_2 = a_3$. (If not we can recycle so this works.) Now since φ is an automorphism, and it is injective, we cannot have anymore coincidences. That is

$$b_2 \neq b_3, \quad a_2 \neq b_3, \quad \text{and} \quad b_2 \neq a_3. \tag{3}$$

Similarly, since $(1\,2)$ and $(1\,i)$ do not commute for every other integer $i > 3$, the permutation $(a_i\ b_i)$ also cannot be disjoint from the permutation $(a_2\ b_2)$. This time, possibly after recycling $(a_i\ b_i)$, we can assume that a_i equals a_2 or b_2. I prove that it cannot be b_2, and in so doing, prove that $a_2 = a_3$ (hence proving that $a_i = a_2$ for $i = 2, \ldots n$).

Claim. For all $i > 3$, $b_2 \neq a_i$

Proof of Claim. Assume otherwise, that is assume $b_2 = a_i$ for the integer i. Since $(1\ 3)$ also does not commute with $(1\ i)$, we also must have that their images under φ, which are $\varphi((1\ 3)) = (a_2\ b_3)$ and $\varphi((1\ i)) = (b_2\ b_i)$, cannot be disjoint. By (3), we have that both letters involved in the first, a_2 and b_3, do not equal b_2. Moreover, if $a_2 = b_i$, then

$$\varphi((1\ i)) = (b_2\ a_2) = (a_2\ b_2) = \varphi((1\ 2))$$

which is impossible since φ is injective and $(1\ 2) \neq (1\ i)$.

Therefore, since it's the only remaining possibility, we must have that $b_i = b_3$. Hence

$$\varphi((1\ 2)) = (a_2\ b_2) \qquad \varphi((1\ 3)) = (a_2\ b_3) \quad \text{and} \quad \varphi((1\ i)) = (b_2\ b_3)\,.$$

Now we calculate before and after applying φ:

- (**Before**)

$$\begin{aligned}(1\ 2)\circ(1\ 3)\circ(1\ i) &= (2\ 1\ 3)\circ(1\ i)\\ &= (3\ 2\ 1)\circ(1\ i)\\ &= (3\ 2\ 1\ i)\\ &= (1\ i\ 3\ 2)\,.\end{aligned}$$

- (**After**)

$$\begin{aligned}\varphi((1\ 2)\circ(1\ 3)\circ(1\ i)) &= \varphi((1\ 2))\circ\varphi((1\ 3))\circ\varphi((1\ i))\\ &= (a_2\ b_2)\circ(a_2\ b_3)\circ(b_2\ b_3)\\ &= (b_2\ a_2\ b_3)\circ(b_2\ b_3)\\ &= (a_2\ b_3\ b_2)\circ(b_2\ b_3)\\ &= (a_2\ b_3)\circ\cancel{(b_2\ b_3)\circ(b_2\ b_3)}\\ &= (a_2\ b_3)\,.\end{aligned}$$

But we have our long awaited contradiction since in the before calculation we ended up with a 4–cycle (here, we're using the assumption that $i > 3$), hence an element of order 4, but in the after calculation we ended up with a 2–cycle, whence an element of order 2. But as an isomorphism, φ would have to take an element of order 4 to an element of order 4, which didn't happen. Thus we finally have our contradiction. ∎

Therefore, for each $i = 2, 3, \ldots, n$, we have $a_i = a_2$. Therefore, for all $2 \leq i \leq n$ we have just proven that

$$\varphi((1\ i)) = (a_2\ b_i)\,. \tag{4}$$

Among other things, since φ is a bijection, this proves, that all of the b_i's are distinct and that none of them are equal to a_1. So, if we denote σ the map which takes $i \mapsto b_i$ for $i \geq 2$ and takes $1 \mapsto a_2$, then σ is an element of $\mathcal{S}_n$ (i.e. it's a bijection).

Claim. The maps φ and c_σ do the same thing to every $(1\ i)$.

Proof of Claim. Since conjugation in $\mathcal{S}_n$ is replacement, we have that

$$c_\sigma(1\ i) = (\sigma(1)\ \sigma(i)) = (a_2\ b_i).$$

But (4) says that φ does the same exact thing to $(1\ i)$. ■

Therefore, (2) has been satisfied and φ is an inner automorphism (moreover $\varphi = c_\sigma$). ■

Here is an example of how this is useful (the same basic proof applies to $\mathcal{S}_n$ for all $n \neq 6$. We will return to this in Exercise 10.63 when we get better counting methods):

Corollary 9.51. *All automorphisms of $\mathcal{S}_4$ are inner.*

Proof. An automorphism must take elements of order 2 to elements of order 2 and take conjugacy classes to conjugacy classes. The only elements of order 2 in $\mathcal{S}_4$ are the 3 $\underline{2}\,\underline{2}$–cycles and the 6 $\underline{2}$–cycles. Since the number 3 does not equal the number 6 (deep, I know) no automorphism can swap these two cycle types (conjugacy classes). Therefore every automorphism of $\mathcal{S}_4$ must take $\underline{2}$–cycles to $\underline{2}$–cycles. Therefore, by Theorem 9.50 every automorphism of $\mathcal{S}_4$ must be an inner automorphism. ■

The idea of the more general proof is, like in this proof, to show that the number of $\underline{2}$–cycles does not match the number of elements of any other cycle type which consist of elements of order 2. This is only false in $\mathcal{S}_6$ where there are a total of 15 $\underline{2}$–cycles, but there are also a total of 15 $\underline{2}\,\underline{2}\,\underline{2}$–cycles! Indeed, there is an automorphism (which cannot be inner since it's swapping around cycle types in a non-trivial way) of $\mathcal{S}_6$ which swap these 15 $\underline{2}\,\underline{2}\,\underline{2}$–cycles with the 15 $\underline{2}$–cycles. Very strange indeed!

9.7 Generators of Binary Groups

In this section we consider the binary groups first introduced on 387. For $X \subseteq \mathbb{R}^3$ this was defined as follows:

$$B(X) \stackrel{\text{def}}{=} \{e^{\frac{\theta}{2}\widehat{\mathbf{u}}} : \varrho_{\theta,\widehat{\mathbf{u}}} \text{ is a symmetry of } X\}$$

or written in a more fancy way as

$$\mathbf{R}^{-1}(S(X))$$

where $S(X)$ is the symmetry group of X and $\mathbf{R}$ is the map

$$\begin{aligned}\mathbf{R}&: \mathbb{S}^3 \mapsto \mathbf{SO}_3(\mathbb{R})\\ &: e^{\frac{\theta}{2}\widehat{\mathbf{u}}} \mapsto \varrho_{\theta,\widehat{\mathbf{u}}}\end{aligned}$$

discussed at length in Section C.7 which begins on page 779. Indeed, define the relation

$$\mathbf{S}: [0, 2\cdot\pi] \times \mathbb{S}^2 \to \mathbb{S}^3: \qquad (\theta, \widehat{\mathbf{u}}) \mapsto e^{\frac{\theta}{2}\widehat{\mathbf{u}}}$$

For the rest of the section, we assume $S(X)$ is finite.

For instance, this is the case for X any sort of polyhedra (since symmetries of X are determined by what they do to the finite vertex set of X. (However it is not true for $X = \mathbb{S}^2$ or $X = \mathbb{S}^1$.)

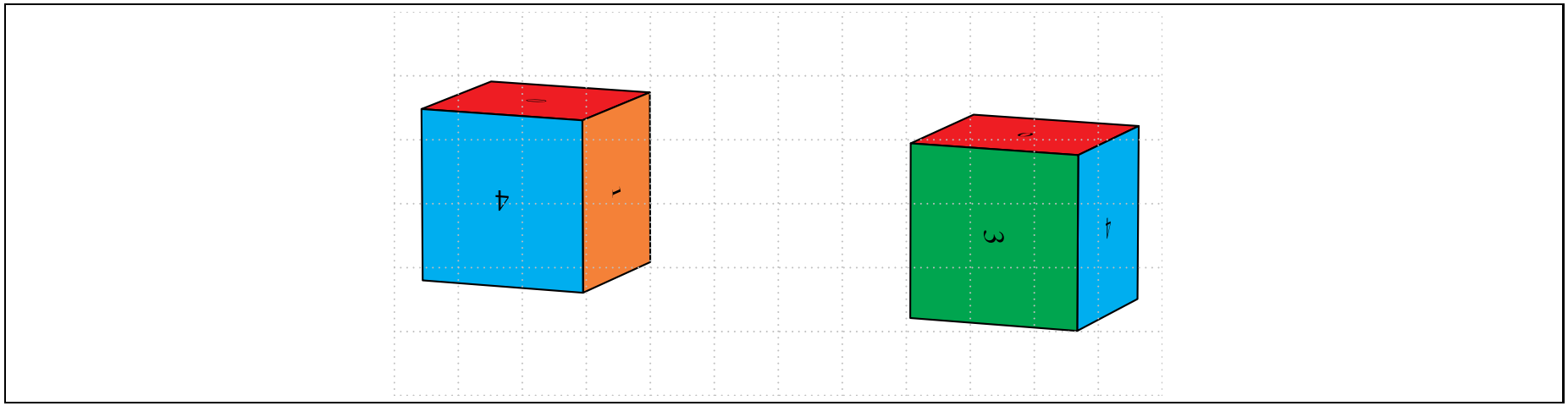


Figure 9.2: The effect of $\varrho_{\frac{\pi}{2},\mathbf{k}}$ on the cube.

For example, since rotation along the upper face of a cube by $\frac{\pi}{2}$ (i.e. by a 90° angle) is a symmetry of a cube (given by $\varrho_{\frac{\pi}{2},\mathbf{k}}$, we have that (Thought of as a subset of $\mathbb{R}^3$ in the usual way. See Figure 9.2)

$$e^{\frac{\pi}{4}\mathbf{k}} = \cos\frac{\pi}{4} + \sin\frac{\pi}{4}\,\mathbf{k} = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{k} \in B\left(\blacklozenge^3\right).$$

Of course since

$$\varrho_{\frac{\pi}{2},\mathbf{k}} = \varrho_{\frac{3\pi}{2},-\mathbf{k}} \Rightarrow \cos\frac{3\pi}{4} + \sin\frac{3\pi}{4}\,(-\mathbf{k}) = -\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\mathbf{k} \in B\left(\blacklozenge^3\right)$$

as well.[I] We discuss this more at length in Section C.42 coming up on page 783.

First a general proposition which is of independent interest. Roughly it says that the generators of the kernel and image[II] of a homomorphism generate the group.

PROPOSITION 9.52. *Let $\varphi\colon G \to H$ be a homomorphism of groups. Suppose that $\mathcal{I}, \mathcal{K} \subseteq G$ have the properties that*

(1) *$\varphi(\mathcal{I})$ generates $\mathrm{Im}(\varphi)$*

(2) *$\mathcal{K}$ generates $\mathrm{Ker}(\varphi)$,*

then $\langle \mathcal{I} \cup \mathcal{K} \rangle = G$.

[I] This reflects the fact that we always have that

$$\mathsf{q} \in B(X) \Leftrightarrow -\mathsf{q} \in B(X)$$

which we could and have proved many different ways. One proof along the lines just presented is

$$e^{\frac{2\pi-\theta}{2}(-\widehat{\mathbf{u}})} = e^{\left(\frac{\theta}{2}-\pi\right)\widehat{\mathbf{u}}} = e^{\frac{\theta}{2}\widehat{\mathbf{u}}} * e^{-\pi\widehat{\mathbf{u}}} = -e^{\frac{\theta}{2}\widehat{\mathbf{u}}}$$

and noting that

$$\varrho_{\theta,\widehat{\mathbf{u}}} = \varrho_{2\pi-\theta,-\widehat{\mathbf{u}}}.$$

[II] This is sloppy, I really mean a section of a generating set.

Proof. Let $g \in G$. We need to show that g can be written as a product of elements of

$$\mathcal{I} \quad \text{and} \quad \mathcal{K}.$$

Indeed, $\varphi(g) \in \text{Im}(\varphi)$ so there exists

$$\exists y \in \langle \mathcal{I} \rangle \text{ s.t. } \varphi(g) = \varphi(y).$$

But this implies that $\varphi(gy^{-1}) = e_H$, i.e. $gy^{-1} \in \text{Ker}(\varphi)$. But by assumption, $\mathcal{K}$ generates this kernel, therefore $gy^{-1} \in \langle \mathcal{K} \rangle$ and there exists a $z \in \langle \mathcal{K} \rangle$ so that

$$gy^{-1} = z \Rightarrow g = zy. \tag{5}$$

This completes the proof. ∎

Actually we proved the stronger result:

Corollary 9.53. *In the notation of the proposition we have*

$$G = \langle \mathcal{K} \rangle \cdot \langle \mathcal{I} \rangle.$$

Proof. This is exactly what we showed in (5). ∎

Rant 9.54. If you have taken linear algebra, you may notice this proof is very similar to the proof of the so called fundamental theorem of Linear Algebra which says that for any vector space V and linear transformation, T, with domain V

$$V \cong \text{Im}(\mathbf{A}) \times \text{Ker}(\mathbf{A}).$$

As with most things, the result for groups is not quite as powerful as it for vector spaces. The reason, in this case, comes down to the fact that we cannot necessarily pick $\mathcal{I}$ and $\mathcal{K}$ to generate disjoint subgroups of G like we always can in linear algebra. For example, for the epimorphism

$$\begin{aligned} \mathbf{R}&: Q_8 \to \mathbb{Z}_2 \times \mathbb{Z}_2 \\ &: \mathbf{i} \mapsto ([1]_2, [0]_2) \\ &: \mathbf{j} \mapsto ([0]_2, [1]_2) \end{aligned}$$

we have that the kernel of $\mathbf{R}$ is $\{\pm 1\}$. However, no matter what we take for the set $\mathcal{I}$ (in the notation of the proposition) we must have that $\langle \mathcal{I} \rangle$ is all of Q_8. For instance, minimal possible choices for $\mathcal{I}$ include $\{\mathbf{i}, \mathbf{j}\}$, $\{-\mathbf{i}, \mathbf{j}\}$, $\{\mathbf{i}, \mathbf{k}\}$, $\{\mathbf{j}, \mathbf{k}\}$, etc., all of which generate all of Q_8. This problem is related to the fact that Q_8 contains no subgroup isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

Arguably (very!) this distinction between vector spaces and groups is the most distinctive one, from a (category) theoretic point of view.

Since we always have the kernel of the restricted map

$$B(X) \mapsto S(X)$$

is ± 1 the proposition guarantees us that that if $\mathbf{R}(\mathsf{q})$ and $\mathbf{R}(\mathsf{p})$ generate $S(X)$, then

$$\{\mathsf{q}, \mathsf{p}, -1\}$$

generate $B(X)$.

We now show that if we are careful picking q and p there's no need for the -1. Indeed, if either q or p have even order, then since -1 is the only element of order 2 in $\mathbb{S}^3$, some power of p or q is equal to -1, and in this case, $-1 \in \langle \mathsf{q}, \mathsf{p} \rangle$, and so $\langle \mathsf{q}, \mathsf{p}, -1 \rangle = \langle \mathsf{q}, \mathsf{p} \rangle$.

THEOREM 9.55. *If* q *and* p *are quaternions whose corresponding rotations generate* $S(X)$ *and at least one of which have even order, then*

$$B(X) = \langle \mathsf{q}, \mathsf{p} \rangle .$$

■

Now we show that we can always pick things in such a way that either p or q has even order. (We can actually pick things so that both have even order.)

NOTATION 9.56. Given a rotation, ρ, we denote the two quaternions which map to ρ under $\mathbf{R}$ as ${}^{-1}\rho$.

In other words, ${}^{-1}\rho$ is simply the inverse image of ρ under $\mathbf{R}$, i.e. $\mathbf{R}^{-1}(\rho)$ (our notation is just to save space.) Therefore, if $\rho = \varrho_{\theta,\widehat{\mathbf{u}}}$, then ${}^{-1}\rho$ is

$$\left\{\cos\frac{\theta}{2} + \sin\frac{\theta}{2}\,\widehat{\mathbf{u}}, -\cos\frac{\theta}{2} - \sin\frac{\theta}{2}\,\widehat{\mathbf{u}}\right\}.$$

LEMMA 9.57. *Let* $\mathsf{q} \in \mathbb{S}^3$ *which has finite order. Then,*

(1) -1 *is a power of either* q *or* $-\mathsf{q}$. [III]

(2) *either* q *or* $-\mathsf{q}$ *has even order.* [IV]

PROOF. Both statements are clearly really the same statement (if you are not convinced of that now, you should be by the time we finish the proof). There are two cases based on the parity of $|\mathsf{q}|$.

$|\mathsf{q}|$ even Suppose q has even order. Then there is an integer k so that $|\mathsf{q}| = 2k$. Hence q^k has order 2. But the only quaternion of order 2 is -1, and so we are done.

$|\mathsf{q}|$ odd Suppose q has odd order, $2k+1$. Then

$$(-\mathsf{q})^{2k+1} = (-1)^{2k+1}\mathsf{q}^{2k+1} = (-1)\cdot(1) = -1.$$

and we have proven that -1 is a power of $-\mathsf{q}$.

■

Since we have

$$\mathbf{R}_{\mathsf{q}} = \mathbf{R}_{-\mathsf{q}}$$

for every symmetry of X we can select a corresponding element of even order in $B(X)$.

THEOREM 9.58. *If* ρ_1 *and* ρ_2 *generate* $S(X)$*, then*

$$\rho_1^+ \quad \textit{and} \quad \rho_2^+$$

generate $B(X)$ *where* ρ_1^+ *and* ρ_2^+ *are a choice of an element of even order in* ${}^{-1}\rho_1$ *and* ${}^{-1}\rho_2$ *resepctively.*

REMARK 9.59. There's no real reason that I'm assuming that $S(X)$ is generated by just two symmetries, the same statement is true no matter how many generators we take for $S(X)$. However, in practice this number always turns out to just be 2.

[IV] I'm not ruling out the possibility that both q and $-\mathsf{q}$ have even order. I'm also not ruling out the possibility that -1 is a power of both.

Proof. So far we know that $B(X)$ is generated by

$$B(X) = \langle \varrho^+,\ \varrho'^+,\ -1 \rangle$$

so we just need to account for the -1. But since -1 is a power of ϱ^+ we know that

$$-1 \in \langle \varrho^+,\ \varrho'^+ \rangle \Rightarrow \langle \varrho^+,\ \varrho'^+ \rangle = \langle \varrho^+,\ \varrho'^+,\ -1 \rangle$$

and we are done. ∎

Remark 9.60. Really, we only need one of the elements picked to have even order since this would already guarantee that we have -1 in the resulting group.

Example 9.61. It's easy to pick many generating sets for $S(X)$ when X is a platonic solid. For instance any pair of rotations of minimal order along adjacent faces or vertices work.

For example, for $X = \blacklozenge^3$ we have that

$$S\left(\blacklozenge^3\right) = \left\langle \overset{\rho_1}{\underset{\shortparallel}{\varrho_{\frac{\pi}{2},\mathbf{i}}}}, \overset{\rho_2}{\underset{\shortparallel}{\varrho_{\frac{\pi}{2},\mathbf{j}}}} \right\rangle.$$

Moreover we have

$$^{-1}(\rho_1) = \left\{ \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{i}, -\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\mathbf{i} \right\} \qquad ^{-1}(\rho_2) = \left\{ \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{j}, -\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\mathbf{k} \right\}.$$

As all elements involved happen to have even order, any combination of one from the first with one from the second will generate $B\left(\blacklozenge^3\right)$. For instance, disliking negative signs,

$$B\left(\blacklozenge^3\right) = \left\langle \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{i}, \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{j} \right\rangle.$$

It's also easy to see that $S(X)$ is always generated by a (minimal order) rotation about a face and a vertex in the that face. For example, for the cube if v the vertex in the upper, front, right-hand corner, $\begin{bmatrix}1\\1\\1\end{bmatrix}$, then

$$S(X) = \left\langle \overset{\rho}{\underset{\shortparallel}{\varrho_{\frac{\pi}{2},\mathbf{k}}}}, \overset{\rho'}{\underset{\shortparallel}{\varrho_{\frac{2\pi}{3},v}}} \right\rangle.$$

We can see this directly as follows: using ρ, we can move any face except for the one with center $-\mathbf{k}$ to a face which contains v. For the pesky face centered at $-\mathbf{k}$ we can first use ρ' to move it to a face other than that pesky one, and then use ρ in the same way we had done originally to end up, in either case, with a face containing v as a vertex. Then using ρ' we can move this new face up to the face centered at $\mathbf{k}$. Therefore, for any symmetry R of a cube, by doing the process just outlined, we have shown there is a combination of ρ's and ρ''s which does the same thing as R does to the face centered at $\mathbf{k}$. But by then, if necessary, applying ρ's to rotated this top face into place, we now see there is a product of ρ's and ρ''s which do exactly the same thing to **every point** of the top face as our arbitrary R does. This clearly means that R is equal to this product (since their quotient does nothing to an entire face which of course means that this quotient is the identity and the two must be equal).

Now using the theorem, we see that

$$^{-1}(\rho) = \left\{ \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{k}, -\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\mathbf{k} \right\} \qquad ^{-1}(\rho') = \left\{ \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}, -\frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k} \right\}.$$

Therefore we also have that (we must take the first listed quaternion for ρ'^{+}, but it doesn't matter which we take for ρ^{+})

$$B\left(\blacklozenge^3\right) = \left\langle \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{k}, \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k} \right\rangle.$$

EXERCISE 9.30. Prove that

$$B\left(\blacktriangleright^3\right) = \left\langle \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}, \mathbf{i} \right\rangle.$$

SOLUTION

Hint! Orient $\blacktriangleright^3$ inside of $\mathbb{R}^3$ in such a way that

$$S\left(\blacktriangleright^3\right) = \left\langle \varrho_{\frac{\pi}{2},\mathbf{i}}, \varrho_{\frac{2\pi}{3},\mathbf{i}+\mathbf{j}+\mathbf{k}} \right\rangle$$

and then use the theorems from this section.

This exercise shows us that the converse of Theorem 9.55 isn't true in the sense that a binary group can be generated by elements of odd order.

EXERCISE 9.31. Show that

$$B\left(\blacktriangleright^3\right) = \left\langle -\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}, -\frac{1}{2} - \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k} \right\rangle.$$

SOLUTION

Note that neither generator has odd order.

In Appendix C, (after aligning $\bigstar^3$ in $\mathbb{R}^3$ in an obvious way) Equation (12) on page 790 shows us that

$$B\left(\bigstar^3\right) = \left\langle \mathbf{i}, \frac{1}{2}\left(\phi + \mathbf{i} + \phi^{-1}\mathbf{j}\right) \right\rangle.$$

In summary, the binary groups of the platonic solids[V] are:

$$B\left(\blacktriangleright^3\right) = \left\langle \mathbf{i}, \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k} \right\rangle$$

$$B\left(\blacklozenge^3\right) = \left\langle \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{i}, \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{j} \right\rangle$$

$$B\left(\bigstar^3\right) = \left\langle \mathbf{i}, \frac{1}{2}\phi + \frac{1}{2}\mathbf{i} + \frac{1}{2}\phi^{-1}\mathbf{j} \right\rangle.$$

9.8 THE ALTERNATING GROUP

Recall that we defined the (working definition for) the set of even permutations, $\mathcal{A}_n$, as the kernel of the homomorphism

$$\varepsilon : \mathcal{S}_n \to \{\pm 1\}$$

[V] We only need three groups since the dodecahedron has the same symmetry group as the icosahedron and the octahedron has the same symmetry groups as the cube.

by first writing an element of $\mathcal{S}_n$ as a permutation matrix and then taking the determinant of this permutation matrix. In less fancy language a permutation is even if its sign is 1 and *odd* if its sign is -1. Everybody has their favorite definition of the sgn, but every one has one of two problems: either it's a bit overwhelming, or it's very difficult to verify that it is a homomorphism or even a function.

Remark 9.62. A common definition is to define a permutation to be even iff it can be written as the product of an even number of transpositions. While less abstract, it is incredibly difficult to show that this function is well defined! For instance, the identity can be written is some crazy ways! For instance, if you have nothing better to do you can verify that

$$(1) = (1\,4)(1\,2)(2\,3)(3\,4)(1\,4)(2\,4)(1\,3)(1\,2)(2\,3)(1\,3)(3\,4)(2\,4).$$

Therefore, it's even kinda surprising that the *identity* **can not** be written as the product of an odd number of transpositions! For instance, we have just written (1) as a product of (the even number) 12 transpositions.

The previous definition that we gave in terms of crossing number is perhaps the fastest and most elementary definition to give. Also, it is clearly well defined. Again, however, it's difficult to prove that the sign turns out to be a homomorphism. Indeed the crossing number certainly isn't a homomorphism in any sense, only its parity is.

Exercise 9.32. Write the identity in 5 different more ways as products of at least 4 different transpositions. Verify in each case that the number of transpositions used is even.

SOLUTION

Amazingly, by the end of this section I hope that you will be convinced that it really doesn't matter how we define this, but the calculations themselves are quite easy! While this definition might seem to be on the abstract side, we illustrate a way that we can completely bypass all of this simply by using the generating set of $\mathcal{S}_n$. It allows us to only calculate ε of a single element directly!

Lemma 9.63. *The permutation* $(1\,2) \in \mathcal{S}_n$ *is odd.*

Proof. The permutation matrix of $(1\,2)$ is the n by n matrix

$$\begin{pmatrix} 0 & 1 & 0 & \cdots 0 \\ 1 & 0 & 0 & \cdots 0 \\ 0 & 0 & 1 & \cdots 0 \\ \vdots & \vdots & \vdots & \cdots \vdots \\ 0 & 0 & 0 & \cdots 1 \end{pmatrix}.$$

which we will denote by Π for the duration of this proof. This is the matrix with 1's down the diagonal except for the upper left hand corner is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Now the top row of the matrix Π has only a single non-zero entry which shows up in the second column and happens to be a 1. While the submatrix one obtains by deleting the first row and second column, which we have been denoting as $\Pi^{(1,2)}$, is simply the $n-1$ by $n-1$ identity matrix, whose determinant is 1. Hence we have

$$\det \Pi = \sum_{i=1}^{1+i} (-1)^i \Pi_{1,i} \cdot \det \Pi^{(1,i)} = (-1)^{1+2=3} \Pi_{1,2} \det \Pi^{(1,2)} = (-1) \cdot (1) \cdot (1) = -1$$

as desired. ∎

Remark 9.64. It's even easier to see this in terms of the original definition of the sign of a permutation in terms of a crossing number. This is because the crossing number of $(1\ 2)$ is obviously just 1.

Again, you guys should notice that from now on we no longer need to talk about permutation matrices at all! The next lemma kicks up the previous lemma a bit. In the previous lemma we showed that a particular 2–cycle was odd, in the next lemma we show that *all* 2–cycles are odd.

Lemma 9.65. *The sign of any 2–cycles is -1. That is all 2–cycles are odd.*

Proof. All 2–cycles are conjugate. That is for any 2–cycle (i.e. transposition) τ, there exists a permutation σ so that

$$\tau = \sigma \circ (1\ 2) \circ \sigma^{-1}.$$

Now ε is a homomorphism into an abelian group. Thus we have

$$\begin{aligned}
\varepsilon\tau &= \varepsilon(\sigma \circ (1\ 2) \circ \sigma^{-1}) \\
&= \varepsilon\sigma \cdot \varepsilon(1\ 2) \cdot \varepsilon\sigma^{-1} && \text{(since } \varepsilon \text{ is a hom)} \\
&= \varepsilon\sigma \cdot \varepsilon(1\,2) \cdot (\varepsilon\sigma)^{-1} && \text{(since } \varepsilon \text{ is a hom)} \\
&= \varepsilon\sigma \cdot (\varepsilon\sigma)^{-1} \cdot \varepsilon(1\,2) && \text{(since } \{-1,+1\} \text{ is abelian)} \\
&= \varepsilon(1\ 2) \\
&= -1 && \text{(by previous lemma).}\blacksquare
\end{aligned}$$

Lemma 9.66. *The sign of an n–cycle is $(-1)^{n-1}$. Therefore an n–cycle is even iff the positive integer n is odd.*

Proof. By repeated use cut-and-paste we can write an n–cycle can be written as a product of $n-1$, 2–cycles as follows:

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

Thus

$$\begin{aligned}
\varepsilon(a_1 \cdots a_n) &= \varepsilon\Big((a_1\,a_2)(a_2\,a_3)\cdots(a_{n-1}\,a_n)\Big) \\
&= \varepsilon(a_1\,a_2)\,\varepsilon(a_2\,a_3)\cdots\varepsilon(a_{n-1}\,a_n) \\
&= (-1)^{n-1}
\end{aligned}$$

as desired. ■

We can combine these Lemmas to make it almost trivial (probably one of the easiest things we do in this course!) to calculate the sign of a permutation.[I]

Example 9.67.

(1) The permutation $(1\ 2)(3\ 4)$ has $\varepsilon((1\ 2)(3\ 4)) = (-1)^2 = 1$. Thus this permutation is even.

[I] It's even easier than calculating the order of a permutation since there we needed to write the permutation as a disjoint union of cycles, and with this we don't

(2) The permutation $(1\,2\,6\,11\,4\,3)(3\,4\,6)(4\,11\,17)$ has

$$\begin{aligned}\varepsilon(1\,2\,6\,11\,4\,3)(3\,4\,6)(4\,11\,17) &= \varepsilon(1\,2\,6\,11\,4\,3)\varepsilon(3\,4\,6)(4\,11\,17)\\ &= (-1)^{6-1}(-1)^{3-1}(-1)^{3-1}\\ &= -1\cdot 1\cdot 1\\ &= -1\end{aligned}$$

and is therefore odd.

THEOREM 9.68. *A permutation is even iff its cycle decomposition has an even number of cycles of even length.*

PROOF. Write a permutation

$$\sigma = \sigma_1\cdots\sigma_k\cdot\tau_1\cdots\tau_m.$$

where each σ_i has even length and each τ_i has odd length. Therefore the integer k is the number of cycles in the cycle decomposition of σ which have even length.

Then since ε is o.p. we have that

$$\varepsilon(\sigma) = \varepsilon(\sigma_1)\cdots\varepsilon(\sigma_k)\cdot\varepsilon(\tau_1)\cdots\varepsilon(\tau_m) = (-1)^k\cdot(+1)^m = (-1)^k.$$

As such, σ is even iff k is even, i.e. iff the number of cycles of even length is even. ■

THEOREM 9.69. *The following table holds* [II]

$\circ$	Even	Odd
Even	Even	Odd
Odd	Odd	Even

PROOF. This is the same table as the group $\{\{2\}\} = \{\pm 1\}$ (whose operation is the usual multiplication of real numbers). Let's do this out (e=even, o=odd)

$$\begin{aligned}
&\varepsilon(\sigma) = +1, \varepsilon(\tau) = +1 && \Rightarrow \varepsilon\sigma\tau = (+1)\cdot(+1) = +1 && \Rightarrow \sigma \text{ e}, \tau \text{ e} && \Rightarrow \sigma\circ\tau \text{ is e}^{\text{III}}\\
&\varepsilon(\sigma) = +1, \varepsilon(\tau) = -1 && \Rightarrow \varepsilon(\sigma\tau) = (+1)\cdot(-1) = -1 && \Rightarrow \sigma \text{ e}, \tau \text{ o} && \Rightarrow \sigma\circ\tau \text{ is o}\\
&\varepsilon(\sigma) = -1, \varepsilon(\tau) = +1 && \Rightarrow \varepsilon(\sigma\tau) = (-1)\cdot(+1) = -1 && \Rightarrow \sigma \text{ o}, \tau \text{ e} && \Rightarrow \sigma\circ\tau \text{ is o}\\
&\varepsilon(\sigma) = -1, \varepsilon(\tau) = -1 && \Rightarrow \varepsilon(\sigma\tau) = (-1)\cdot(-1) = +1 && \Rightarrow \sigma \text{ o}, \tau \text{ o} && \Rightarrow \sigma\circ\tau \text{ is e}
\end{aligned}$$

Which are all the cases in the table. ■

In Theorem 10.107 we will prove the general formula

$$|G| = |\operatorname{Im}(\varphi)|\cdot|\operatorname{Ker}(\varphi)|$$

for any group homomorphism φ with domain G. In our case ε is a group homomorphism with $|\operatorname{Ker}(\varepsilon)| = |\mathcal{A}_n|$ and $|\operatorname{Im}(\varphi)| = |\mathbb{Z}_2| = 2$. Thus we have that

$$|\mathcal{A}_n| = \frac{\mathcal{S}_n}{2} = \frac{n!}{2}.$$

Before proving this general theorem, however, we give a direct proof of this now. It may help indicate why the general theorem, Theorem 10.107, might be true.

[II] Written out, this table says that any even permutation composed with an even permutation is even, even composed with odd is odd, odd composed with even is odd and odd composed with odd is even.

THEOREM 9.70. *For $n > 1$, in $\mathcal{S}_n$, half of all permutations are even and half are odd. That is*

$$|\mathcal{A}_n| = \frac{n!}{2}.$$

PROOF. As usual, we denote $\mathcal{A}_n$ the set of even permutations, and for this proof we denote the set of odd permutations in $\mathcal{S}_n$ as $\mathcal{O}_n$. Since every permutation is either even or odd (and none are both) we have that

$$\#\mathcal{S}_n = \#\mathcal{A}_n + \#\mathcal{O}_n.$$

So, in order to establish the theorem, we need to prove that $\#\mathcal{A}_n = \#\mathcal{O}_n$. We do this by forming an explicit bijection between the two.

Since we are assuming that $n > 1$, $\mathcal{S}_n$ contains at least one odd permutation, $\mathbf{o}$.[IV] Since $\mathbf{o}$ is odd, the previous theorem, Theorem 9.69, says that for any $\sigma \in \mathcal{S}_n$, the permutations σ and $\mathbf{o} \circ \sigma$ have opposite parity. In other words, one is even iff the other is odd. Said in a fancier way, for any odd permutation $\mathbf{o}$, left multiplication by $\mathbf{o}$ restricts to a well defined functions

$$\begin{aligned} \mathrm{L}_{\mathbf{o}}|_{\mathcal{A}_n} &: \mathcal{A}_n \to \mathcal{O}_n \\ &: \sigma \mapsto \mathbf{o} \circ \sigma \end{aligned} \qquad \begin{aligned} \mathrm{L}_{\mathbf{o}}|_{\mathcal{O}_n} &: \mathcal{O}_n \to \mathcal{A}_n \\ &: \sigma \mapsto \mathbf{o} \circ \sigma. \end{aligned}$$

CLAIM. $\mathrm{L}_{\mathbf{o}}|_{\mathcal{A}_n}$ is a bijection with inverse $\mathrm{L}_{\mathbf{o}^{-1}}|_{\mathcal{O}_n}$.[V]

PROOF OF CLAIM. This is a direct check:

$$\begin{aligned} \left(\mathrm{L}_{\mathbf{o}}|_{\mathcal{A}_n} \circ \mathrm{L}_{\mathbf{o}^{-1}}|_{\mathcal{O}_n}\right)(\sigma) &= \mathrm{L}_{\mathbf{o}}|_{\mathcal{A}_n}\left(\mathrm{L}_{\mathbf{o}^{-1}}|_{\mathcal{O}_n}(\sigma)\right) \\ &= \mathrm{L}_{\mathbf{o}}|_{\mathcal{A}_n}\left(\mathbf{o}^{-1} \circ \sigma\right) \\ &= \mathbf{o} \circ (\mathbf{o}^{-1} \circ \sigma) \\ &= (\mathbf{o}\mathbf{o}^{-1}) \circ \sigma \\ &= \sigma \end{aligned}$$

$$\Rightarrow \mathrm{L}_{\mathbf{o}}|_{\mathcal{A}_n} \circ \mathrm{L}_{\mathbf{o}^{-1}}|_{\mathcal{O}_n} = \mathbf{1}_{\mathcal{O}_n}.$$

Similarly, we also have that

$$\mathrm{L}_{\mathbf{o}^{-1}}|_{\mathcal{O}_n} \circ \mathrm{L}_{\mathbf{o}}|_{\mathcal{A}_n} = \mathbf{1}_{\mathcal{A}_n}.$$

Hence the claim is proven. ■

In particular, we have a bijection $\mathcal{A}_n \to \mathcal{O}_n$, and so $\#\mathcal{A}_n = \#\mathcal{O}_n$ as desired. ■

Taking $\mathbf{o} = (1\ 2)$, the "picture" for this proof would have you line up all the even permutations on one side and all odd permutations on the other. Then connect an even permutation to an odd one if you can left multiply by $(1\ 2)$ to get between the two. See Figure 9.3 for the case $n = 3$.

[IV] Indeed, since $n > 1$, we have $(1\ 2) \in \mathcal{S}_n$. In fact, in this proof everywhere you see an $\mathbf{o}$ you can feel free to replace it with $(1\ 2)$. We use $\mathbf{o}$ instead of $(1\ 2)$ to help drill in the fact that this proof does not depend on any specific properties of the particular permutation, $(1\ 2)$. For instance, it does not depend on the fact $(1\ 2)$ has order 2.

[V] Note that this statement makes sense since $\mathbf{o}^{-1}$ is an odd permutation whenever $\mathbf{o}$ is.

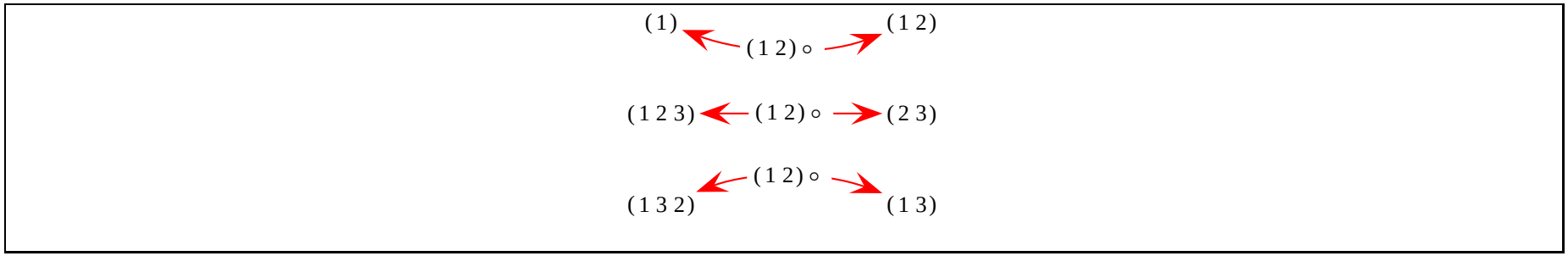


Figure 9.3: Figure for Theorem 9.70. Here, we lined the even permutations up on the left and the odd permutations up on the right.

EXERCISE 9.33. Which of the following permutations are even, which are odd?

SOLUTION

(1) $(1\,3\,6\,11\,13)(1\,6\,9\,12)(3\,1\,4\,5\,9\,2\,6)$

(2) $(1\,2\,3\,4)(5\,6\,7\,8\,9)(1\,2\,3\,4\,5\,6\,7\,8\,9)$

(3) $(1\,2)(3\,4)(1\,3\,4)(3\,1\,4\,5\,9\,2\,6)$

(4) $(1\,7\,9)(2\,6\,4)(3\,1\,4\,5\,9\,2\,6)$.

(5) $\sigma(1) = 12$ $\sigma(2) = 2$ $\sigma(3) = 3$ $\sigma(4) = 1$
$\sigma(5) = 11$ $\sigma(6) = 9$ $\sigma(7) = 5$ $\sigma(8) = 10$
$\sigma(9) = 6$ $\sigma(10) = 4$ $\sigma(11) = 7$ $\sigma(12) = 8$

(6) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 5 & 6 & 1 & 8 & 7 & 3 & 4 \end{pmatrix}$

EXERCISE 9.34. Without doing out any part of the calculation, how can you tell that

SOLUTION

$$(1\,4\,5\,3\,6\,9)(1\,5\,3)(1\,8\,9\,2\,6\,5)(1\,2\,3\,9\,8\,5)(1\,3\,6\,9)$$

is not a 4–cycle?[VI]

EXERCISE 9.35. Show that a permutation is even iff it can be written as the product of an even number of 2-cycles.

SOLUTION

[VI] So this product cannot possible equal $(1\,4\,5\,6)$ or $(1\,3\,2\,6)$, etc.

EXERCISE 9.36. Prove that every odd permutation has even order. Is the converse also true (i.e. are there even permutations with even order)?
Hint! There are at least two proofs of this. The less elegant, but more commonly done proof (for undergraduates) involves using the fact that all odd permutations which happen to be cycles have even length. Then proving the statement for a general permutation by using properties of least common multiples.
The better proof simply uses the fact that for odd σ, $\varepsilon(\sigma^n) = (-1)^n$.
SOLUTION

EXERCISE 9.37. Show that $\mathcal{A}_3 \cong \mathbb{Z}_3$.
Hint! Theorem 8.61.
SOLUTION

EXERCISE 9.38. For $n > 3$, prove that $\mathcal{A}_n$ is centerless.
Hint! There are two cases based on the parity of the number n. For n even, all $n-1$–cycles are even. If n is odd, then all $n-3$ 2–cycles are even. Then use Theorem 7.29 back on page 355.
SOLUTION

EXERCISE 9.39.

(1) Show that in $\mathcal{A}_4$ that all 3-cycles are not conjugate (this is despite the fact that they are all conjugate in the larger group $\mathcal{S}_4$!)

(2) Show that for $n > 4$ that all 3-cycles are conjugate.
Hint! In these cases there is always a two cycle disjoint from any 3-cycle.
SOLUTION

EXERCISE 9.40. In this problem we show that the set of 3–cycles generate $\mathcal{A}_n$.

(1) Show that the product of 2 transpositions is a product of 3-cycles (break this into two cases based on whether or not the two cycles are disjoint).

(2) Use this and Exercise 9.35 just above to show that this implies that the set of 3-cycles generate $\mathcal{A}_n$.
SOLUTION

EXERCISE 9.41. Prove that for any 3–cycle σ, $\lfloor \sigma \rceil_{\mathcal{A}_n} = \mathcal{A}_n$.
Hint! Follow the same logic as was used in Example/Theorem 9.14 with a couple of minor adjustments for $n < 5$.
SOLUTION

In this problem I will allow you to use the following fact (proven in Theorem 10.61): if $H \leqslant G$ and $2 \mid H \mid = \mid G \mid$, then $H \trianglelefteq G$.

Exercise 9.42. Show that $\mathcal{A}_4$ contains no subgroup of order 6.
Hint! Count the number of $\underline{3}$-cycles in $\mathcal{A}_4$!

Solution

In this problem you show that $\mathcal{S}_n$ is isomorphic to a subgroup of $\mathcal{A}_{n+2}$.

Exercise 9.43. For $n > 3$, define

$$\begin{aligned} F &: \mathcal{S}_{n-2} \to \mathcal{S}_n \\ &: \sigma \mapsto \begin{cases} \sigma \circ ((n-1)\ n) & \text{if } \sigma \text{ is odd} \\ \sigma & \text{if } \sigma \text{ is even} \end{cases} \end{aligned}$$ [VII]

Solution

(1) Show that F has image contained in $\mathcal{A}_n$. (That is show that $F(\sigma)$ is even for all $\sigma \in \mathcal{S}_n$)

(2) Show that F is a monomorphism. (So you must show the F is both a homomorphism **and** that F is injective.)

(3) Conclude that $\mathcal{S}_n \cong F(\mathcal{S}_n)$ and so $\mathcal{S}_n$ is (isomorphic to) a $\leqslant \mathcal{A}_{n+2}$.

[VII] A fancy way to write F is as

$$F(\sigma) = \sigma \circ (n-1\ n)^{\frac{1-\varepsilon(\sigma)}{2}}.$$

10 Group Actions and Lagrange's Thm

Groups are the verbs of mathematics. Their primary function is to mix things up, and in the process exploit symmetry. So far we have not come up with a systematic way of creating group homomorphisms (except by conjugating).

10.1 Basic Definitions

Definition 10.1. A group action of a group G on a set X is a function

$$\bullet\colon G \times X \to X^{\text{I}}$$

which satisfies the following two axioms (for all $g, h \in G$ and $x \in X$):

$e_G \bullet x = x$	(Identity acts as identity)	**A1**
$g \bullet (h \bullet x) = gh \bullet x$	(Associativity for actions)	**A2**

We write this more succinctly as $\bullet\colon G \circlearrowright X$ or simply $G \circlearrowright X$. This is read "G hit X." In this situation we say that X is a *G-set* and that G acts on X (via $\bullet$). We call G the *verb* of the action and X the *noun* of the action. The following theorem shows that we could have given a very different looking definition of a group action.

Definition 10.2.

(1) Given a group action $\bullet\colon G \circlearrowright X$, we define a function from G to $\mathcal{F}_X$ by "fixing a variable", that is $\rho^\bullet$ sends g to the function which takes x to $g \bullet x$. Or written out, $\rho^\bullet_g(x) = g \bullet x$. [II] This function is known as the *permutation representation of* $G \circlearrowright X$.

(2) Given a group homomorphism $\rho\colon G \to \mathcal{S}_X$, we define

$$\begin{aligned} \bullet_\rho &\colon G \times X \to X \\ &\colon (g, x) \mapsto \rho_g(x). \end{aligned}$$

THEOREM 10.3. *The image of a permutation representation of an action, $\bullet\colon G \circlearrowright X$, consists of permutations.*[III] *and defines a group homomorphism*

$$\rho^{\bullet}\colon G \to \mathcal{S}_X.$$

On the other hand, if we are given a group homomorphism

$$\rho\colon G \to \mathcal{S}_X,$$

then $\bullet_\rho$ defines an action of G on X. [IV]

REMARK 10.4. Before proving this, we should note that these two processes are inverse to each other since

$$\rho_g^{\bullet_\rho}(x) = g \bullet_\rho x = \rho_g(x)$$

and so $\rho^{\bullet_\rho} = \rho$. At the same time

$$g \bullet_{\rho^\bullet} x = \rho_g^{\bullet}(x) = g \bullet x$$

and so $\bullet_{\rho^\bullet} = \bullet$.

PROOF.

(1) We first show that $\rho^{\bullet}$ satisfies **(A1)** and **(A2)**.

(A1) Since ρ is by definition a group homomorphism $\rho\colon G \to \mathcal{S}_X$ and, as such, it must take the identity of the first group, $e = e_G \in G$ to the identity of the second group $\mathbf{1}_X \in \mathcal{S}_X$, the identity function. Hence $\rho_e = \mathbf{1}_X$. Now if we apply both sides of this equation to a point $x \in X$ we see that

$$e \bullet_\rho x = \rho_e(x) = x$$

as desired. Therefore the identity acts like the identity.

(A2) Again, the elements of the set $\mathcal{S}_X$ are functions, so the operation in this group is function composition. Thus, the fact that

$$\rho : G \to \mathcal{S}_X$$

is a group homomorphism means that $\rho_{g*h} = \rho_g \circ \rho_h$, where $*$ is the operation in G. If we apply both sides of this equation to $x \in X$ we see that

$$\rho_{g*h}(x) \quad \text{and} \quad \left(\rho_g \circ \rho_h\right)(x)$$

are equal for all $x \in X$. But the first equation is $(g * h) \bullet_\rho x$ and the second is $g \bullet_\rho (h \bullet_\rho x)$ which is proves that $\bullet_\rho$ is associative.

In conclusion, we have proved that $\bullet_\rho$ is a group action whenever $\rho\colon G \to \mathcal{S}_X$ is a group homomorphism.

(2) Now we suppose that $\bullet : G \circlearrowright X$ is a group action, i.e. it satisfies **(A1)** and **(A2)**. We need to show that $\rho^{\bullet}$ is a group homomorphism $G \to \mathcal{S}_X$. Obviously we need to show that $\rho^{\bullet}$ is operation preserving, but we can't forget to show that it actually has images that are invertible.[V]

[III] In other words, for every $g \in G$ the function $\rho_g^{\bullet}$ is a bijection.

[IV] See remark after the theorem

[V] Note that for a fixed $x_0 \in X$ that $g \bullet x := x_0$ satisfies **(A2)** but not **(A1)** (if X has more than one point). However, in this case the function $\rho^{\bullet}$ would be operation preserving but would not have images that are bijections!

Let's tackle the first objective first. To this end we consider arbitrary points $g, h \in G$. As we are trying to show the equality of two functions, $\rho_g^\bullet \circ \rho_h^\bullet$ and $\rho_{g*h}^\bullet$ we just need to show that both of these dudes do exactly the same thing to every point x in their domains, which is X.

$$\begin{aligned}
\rho_{g*h}^\bullet(x) &= (g * h) \bullet x && \text{(By def'n of } \rho^\bullet\text{)} \\
&= g \bullet (h \bullet x) && \text{(By (A2))} \\
&= \rho_g^\bullet(h \bullet x) && \text{(By def'n of } \rho^\bullet\text{)} \\
&= \rho_g^\bullet(\rho_h^\bullet(x)) && \text{(By def'n of } \rho^\bullet\text{)} \\
&= (\rho_g^\bullet \circ \rho_h^\bullet)(x) && \text{(By def'n of compositions of functions)}
\end{aligned}$$

as desired.

The last thing we need to do is show that $\rho_g^\bullet$ is a bijection for every $g \in G$. We do this by finding an explicit inverse to it which, as we would expect, winds up being $\rho_{g^{-1}}^\bullet$. For starters, note that for the identity $e_G = e \in G$ we have

$$\rho_e^\bullet(x) = e \bullet x = x \Rightarrow \rho_e^\bullet = \mathbf{1}_X. \tag{1}$$

Using this we see

$$\begin{aligned}
\rho_g^\bullet \circ \rho_{g^{-1}}^\bullet &= \rho_{g*g^{-1}}^\bullet && \text{(Since } \rho^\bullet \text{ is o.p)} \\
&= \rho_e^\bullet && \text{(Since } g * g^{-1} = e_G\text{)} \\
&= \mathbf{1}_X. && \text{(By (1))}
\end{aligned}$$

And the proof that

$$\rho_{g^{-1}}^\bullet \circ \rho_g^\bullet = \mathbf{1}_X$$

now holds by replacing g by g^{-1} and g^{-1} by $g = (g^{-1})^{-1}$. Therefore

$$\left(\rho_g^\bullet\right)^{-1} = \rho_{g^{-1}}^\bullet \Rightarrow \forall g \in G, \rho_g^\bullet \in \mathcal{S}_X.$$

■

Remark 10.5. Careful, we proved that $\rho_g^\bullet$ is a bijection, we certainly didn't prove that $\rho^\bullet$ is a bijection. For instance, for every real $r \neq 0$, we have that $m_r : \mathbb{R} \to \mathbb{R}$ given by multiplication by r is a bijection (this defines a group action of $\mathbb{R}^\times \circlearrowright \mathbb{R}$). However the function $(r, s) \mapsto r \cdot s$ is **not** even close to a bijection. For instance $(1, 0) \mapsto 0$ and so does $(2, 0)$. (Actually for any real number $r \neq 0$, $(r, 0) \mapsto 0$, so the function has an uncountable number of points mapping to 0.)

This gives us two ways of thinking of a group action. The first way, $g \bullet x$, can be thought of as a kind of multiplication. However, just like scalar multiplication of a number and a vector, the things you are actually "multiplying" are very different. The second way, $\rho_g(x)$, can be thought of a way of turning a group into bijections of the noun, X. The advantage of the second version is that it gives us more theoretical insight into what exactly an action is doing: it tells us that an action is a way of seeing a group as "mixing" a set X up. On the other hand, the second definition is typically easier and more straightforward to actually check ... one simply need to show that the two axioms hold. Like many things in life, we are tempted to pick favorites, but in reality you should try to imagine both definitions because they each add important insight into what an action really is.

Remark 10.6. As pointed out in the appendix on Cayley's theorem, Appendix B back on page 753, you should think of functions $X \times Y \to Z$ and $X \to \mathcal{F}\text{un}(Y, Z)$ as being essentially the same thing. The first case should be thought of as functions on two variables, the second case should be thought of getting a parametrized family of functions by fixing the first variable. For instance, in the most familiar case of $X, Y, Z = \mathbb{R}$ we could think of the function $\sin xy$ as a function of two variables, x and y or as a different function for every fixed x, i.e. fixing a variable defines a function of the other variable. For instance, $x = 17$ corresponds to the function $\sin 17y$, $x = \pi$ corresponds to the function $\sin \pi y$ etc. That is $\sin xy$ defines a function $\mathbb{R} \to \mathcal{F}\text{un}(\mathbb{R}, \mathbb{R})$ or from $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$ depending on your perspective.

Exercise 10.1. To understand the steps of this proof a little bit better let's generalize it. Let M be a monoid and let G be a group. Show that an operation preserving map

$$\rho : G \to M$$

has image contained in $\mathcal{U}(M)$ iff $\rho(e_G) = e_M$ (this of course came for free in the case that M was a group, but this is no longer the case when M is just a monoid).

Solution

Definition 10.7. Suppose that we have an action $\bullet : G \circlearrowright X$.

(1) We say that an element $g \in G$ *fixes* an element $x \in X$ if $g \bullet x = x$.

(2) Let $x \in X$. The *stabilizer* of x, denoted G_x, is the set of all elements of G which fix x. In symbols

$$G_x \stackrel{\text{def}}{=} \{g \in G : g \bullet x = x\}.$$

(3) Let $g \in G$. The *fixed point set* of g, denoted X^g, is the set of all elements of X which are fixed by g. In symbols

$$X^g \stackrel{\text{def}}{=} \{x \in X : g \bullet x = x\}.$$

(4) The *orbit* of some $x \in X$, denoted as x/G or $G \bullet x$, is the set

$$\{y : \exists\, g \in G \text{ s.t. } g \bullet x = y\} = \{g \bullet x : g \in G\}.$$

(5) The kernel of an action, $G \circlearrowright X$, is the set of $g \in G$ which fix **every** element of X. That is

$$\text{Ker}(\bullet) = \bigcap_{x \in G} G_x.$$

(6) The action is *faithful* if only the identity fixes everybody.
That is if $g \bullet x = x$ for all $x \in X$, then $g = e_G$, i.e

$$\text{Ker}(\bullet) = \{e_G\}.$$

(7) The action is *free* if only the identity fixes somebody.
That is if $g \bullet x = x$ for some $x \in X$, then $g = e_G$, i.e.

$$\forall x \in X, G_x = e_G \quad \text{i.e.} \quad \forall g \in G, X^g \neq \emptyset \Leftrightarrow g = e_G.$$

(8) An action is *transitive* if any two elements of X are in the same orbit. That is for all $x, y \in X$ there exists a $g \in G$ so that $g \bullet x = y$.

Notice that by definition, free actions are faithful. In words, this is because free actions cannot elements which fix any point while faithful actions can't have elements which fix every point. However, the requirement that an action be free is a fairly rare occurrence. The following simple exercise shows that the kernel of an action is the kernel of a group homomorphism (in particular its a normal subgroup of G).

We also extend the definitions of the stabilizer and fixed point set to include more than one element.

SubDefinition 10.8. Suppose that $S \subseteq G$ and $A \subseteq X$, then define

(1) The set of verbs

$$G_A \stackrel{\text{def}}{=} \{g \in G : g \bullet A = A\}$$

is called the *stabilizer* of A. ($g \bullet A \stackrel{\text{def}}{=} \{g \bullet a : a \in A\}$)

(2) If $g \in G_A$, we say that g *leaves* A *invariant.*

(3) The set of nouns

$$X^S \stackrel{\text{def}}{=} \{x \in X : \forall g \in S, g \bullet x = x\}$$

is called the *fixed point set* of S. We call X^G the *global fixed point set of* $G \circlearrowright X$

Of course these reduce to the above definitions in the special cases that S or A contain a single element. Occasionally, we use the more explicit notations Stab(A) and Fix(S) to denote (respectively) G_A and X^S.

Warning! 10.9. While it's obviously true that

$$X^S = \bigcap_{s \in S} X^s,$$

we only have that, in general,

$$\bigcap_{x \in A} G_x \subseteq G_A. \tag{2}$$

The reason is that its possible to fix A as a set without necessarily fixing every point in A. For instance, a 180° rotation fixes the x-axis as a set (hence leaves it invariant), but it only fixes a single point on the x-axis, namely the origin. In this way we'd have

$$180^\circ \text{ rotation } \in \text{ Stab}(x\text{-axis}) \quad \text{and} \quad 180^\circ \text{ rotation } \notin \text{Stab}((1,0)).$$

In the literature, the stabilizer we are studing could be called the *set-wise stabilizer* of A, while the intersection on the LHS of (2) is called the *point-wise stabilizer* of A. For instance for an action $G \circlearrowright X$ we always have $G_X = G$, even for non-trivial actions. Meanwhile the point-wise stabilizer of X is by definition the kernel of the action. We have little use for point-wise stabilizers so we won't bother to invent any sort of notation to describe it. (In fact you can delete the words "point-wise stabilizer" from your brain at this time!

Exercise 10.2. Prove that

$$x \in X^S \Leftrightarrow S \subseteq G_x.$$

Solution

Exercise 10.3. For an action $G \circlearrowright X$ and $S \subseteq G$

(1) If $S' \subseteq S$ prove that $X^S \subseteq X^{S'}$;

(2) Prove that $X^S = X^{\langle S \rangle}$.

Solution

Exercise 10.4. For an action $\bullet\colon G \circlearrowright X$ with permutation representation $\rho^\bullet$, prove that $\operatorname{Ker}(\bullet) = \operatorname{Ker}(\rho^\bullet)$.

Solution

Exercise 10.5. Prove that TFAE

(1) $G \bullet x = \{x\}$;

(2) $x \in X^G$;

(3) $G_x = G$.

Solution

Using the above exercise we can no prove the following:

Theorem 10.10. *The kernel of any action $G \circlearrowright X$ is a normal subgroup of G.*

Proof. The exercise, Exercise 10.4, says that the kernel of the action is the kernel of some homomorphism. But we know that the kernel of any homomorphism is a normal (see Theorem 8.49) subgroup of its domain. ∎

Theorem 10.11. *An action is faithful iff its corresponding permutation is a monomorphism.*[VI]

Proof. By the exercise, Exercise 10.4, for a faithful action, the corresponding permutation representation

$$\rho^\bullet\colon G \to \mathcal{S}_X$$

has a trivial kernel. But we know, like for any homomorphism of groups, that this is equivalent to $\rho^\bullet$ being a monomorphism. ∎

[VI] This means that we can carve out a subgroup of $\mathcal{S}_X$ which is isomorphic to G. Identifying $\mathcal{S}_X$ with $\mathcal{S}_n$ (whenever $\#X = n$) says we can write G as a symmetric group.

10.2 A Quattuordecillion Examples

There are almost too many examples of group actions. In practice each actions gives us new insight into the group G. However knowing which group action to use in which situation is a highly right-brained task. It's almost like learning u–substitution in a basic calculus course. At first you need to be told what to let u equal in your integral, but eventually you develop a decent intuition. However at first, like every student, you end up making the substitution $u = x$ (yes, it's embarrassing, but most students have tried this at least once!).

Example 10.12.

(1) For any group G and any set X the trivial action, defined as $g \bullet x = x$, is an action. (zzzzz)

(2) The group $\mathcal{S}_n$ acts on the set of letters, $\overline{n}$ via

$$\sigma \bullet i := \sigma(i).$$

More generally, if $G \leqslant \mathcal{S}_X$, then the inclusion map

$$G \to \mathcal{S}_X$$

defines a permutation representation of G on X (since the inclusion of any subgroup into the larger group is a homomorphism). This is clearly a faithful action (since the inclusion map is injective). While these actions are always faithful, they are rarely free. For instance if $n > 2$ $\mathcal{S}_n$ contains the 2–cycles $(1\ 2)$ which fixes everything bigger than 2.

(3) We have that

$$\mathcal{D}_4 \circlearrowright \{\text{4 Vertices of the Square}\}$$

via $g \bullet v = g(v)$. This is an action by definition of the composition of functions (which is the operation in $\mathcal{D}_4$). This action is faithful since if a symmetry of a square fixes all vertices (actually any 2 adjacent vertices) then that symmetry must be trivial. Just this once, I will write down the entire table for this action:

$\bullet$	v_1	v_2	v_3	v_4
$\mathbf{1}_\blacklozenge$	v_1	v_2	v_3	v_4
ρ	v_2	v_3	v_4	v_1
ρ^2	v_3	v_4	v_1	v_2
ρ^3	v_4	v_1	v_2	v_3
$\mathbf{r}$	v_1	v_4	v_3	v_2
$\mathbf{r}\rho$	v_4	v_3	v_2	v_1
$\mathbf{r}\rho^2$	v_3	v_2	v_1	v_4
$\mathbf{r}\rho^3$	v_2	v_1	v_4	v_3

Glancing at the table, we see that the stabilizer of, for instance, the stabilizer of v_2 is $\{\mathbf{1}_\blacklozenge, \mathbf{r}\rho^2\}$. This is because if we look down the column labeled v_2, a v_2 shows up only in the rows labeled $\mathbf{1}_\blacklozenge$ and $\mathbf{r}\rho^2$. Meanwhile the fixed point set of $\mathbf{r}$ is $\{v_1, v_3\}$. This is because if we look across the row of $\mathbf{r}$ these are the two entries which are still in their column. The induced representation is given by (on generators)

$$\rho \mapsto \begin{pmatrix} v_1 & v_2 & v_3 & v_4 \\ v_2 & v_3 & v_4 & v_1 \end{pmatrix} \quad \text{and} \quad \mathbf{r} \mapsto \begin{pmatrix} v_1 & v_2 & v_3 & v_4 \\ v_1 & v_4 & v_3 & v_2 \end{pmatrix}$$

which we get simply by copying over the row next to ρ and $\mathbf{r}$ respectively.

(4) Let G be any group. We define an action, $\bullet_R$, of G on itself, called the *cayley Action* via

$$g \bullet_R x := gx. \qquad \text{(Here } gx \text{ is mult in } G!)^{\text{I}}$$

In fact, we can check that this $\bullet_R$ is an action:

- The identity e_G acts by the identity since $e_G \bullet x = e_G x = x$.
- Associativity for actions is simply a restatement of the associativity that holds in any group. Indeed,

$$g \bullet (h \bullet x) = g \bullet (hx) = g(hx) = (gh)x = (gh) \bullet x.$$

This action is a free action (since if $g \bullet x = gx = x$ implies that $g = e_G$ by right cancelation). In particular, the action if faithful. Therefore, taking the corresponding permutation representation (and a dictionary on G) we have proven

THEOREM 10.13 (Cayley's Theorem). *Suppose G has order n. Then G is (isomorphic to) a subgroup of S_n.* ■

EXERCISE 10.6.

SOLUTION

(1) Define $g \bullet x := xg$. Prove that $\bullet$ does not define an action in general (do this by finding a specific group where this $\bullet$ does not define an action).

(2) Instead, define $g \bullet x := xg^{-1}$. Prove that $\bullet$ does define an action.

The next exercise introduces the notation of a right–action and compares it to what we have been calling an action (which could be called a *left* action).

EXERCISE 10.7. Let G be a group and X a set. A *right action* of G on X is a function

SOLUTION

$$\begin{aligned} \star &: X \times G \to X \\ &: (x, g) \mapsto x \bullet g \end{aligned}$$

so that

(1) $\forall x \in X, x \star e_G = x$.

(2) $\forall x \in X, g_1, g_2 \in G, x \star (g_1 g_2) = (x \star g_1) \star g_2$.

In a sense, the study of (left) actions is not much different from the study of right actions, so we will content ourselves to simply study (left) actions. This is illustrated by the following sequence of exercises.

(1) Suppose that $g \bullet x$ is an (left) action, prove that $x \star g := g^{-1} \bullet x$ defines a right action.

(2) Suppose that $x \bullet g$ is a right action, prove that $x \star g := g^{-1} \bullet x$ defines an action.

(5) The Cayley action is one of two actions where both the noun and the verb are the same, G. The second general action of a group on itself (i.e. both the noun and the verb are the group G) is by conjugation. That is we define $g \bullet_C x \stackrel{\text{def}}{=} gxg^{-1}$. To emphasize when I'm talking about G as a noun versus a verb, I turn G bold when talking about it as a noun. This is an action since

- The identity acts as the identity since
$$x \in \mathbf{G} \Rightarrow e_G \bullet_C x = e_G x e_G^{-1} = x.$$
- $\bullet_C$ is associative since
$$\begin{aligned} \forall x \in \mathbf{G}, g, h \in G \; g \bullet_C (h \bullet_C x) &= g(h \bullet_C x)g^{-1} \\ &= g(hxh^{-1})g^{-1} \\ &= (gh)x(gh)^{-1} \quad \text{(Socks and Shoes)} \\ &= gh \bullet_C x. \end{aligned}$$

Unfortunately, in practice both actions–the Cayley action and the conjugation action–will be denoted as $g \bullet x$, so we need to pay attention to the content to know which one.

What do stabilizers look like? Well, for any $x \in \mathbf{G}$, we have that $g \in G_x$ iff
$$g \bullet x = gxg^{-1} = x.$$
That is, under this action
$$G_x = C_G(x),$$
the centralizer of x. Ironically, this action has a lot of self duality to it because the fixed point set of an element in $\mathbf{G}$ is also the centralizer[II] of this element since
$$\begin{aligned} \mathbf{G}^g &= \{x \in \mathbf{G} : g \bullet x = x\} \\ &= \{x \in \mathbf{G} : gxg^{-1} = x\} \\ &= \{x \in \mathbf{G} : gx = xg\} \\ &= C_G(g). \end{aligned}$$

Since the center of G is the intersection of all of the centralizers of G, we also have
$$\text{Ker}(\bullet_C) = \mathcal{C}(G).$$
Moreover, the global fixed point set of $\bullet_C : G \circlearrowright \mathbf{G}$ is also the center of G. That is
$$\mathbf{G}^G = \mathcal{C}(G).$$

What do orbits look like? Well the orbit of x is the set
$$G \bullet_C x = \{g \bullet x : g \in G\} = \{gxg^{-1} : x \in G\},$$
which by its very definition, is the conjugacy class of x in G. That is
$$G \bullet_C x = \mathcal{CC}_G(x).$$
However, unlike they Cayley action which is always free and transitive, for non-trivial groups, the conjugation action is never free or transitive since
$$g \bullet_C g = ggg^{-1} = g \Rightarrow g \in G_g$$
implies the action is not free.

So $\bullet_C$ is faithful iff G is centerless.

[II] All of this is wrapped up in the statement that $gxg^{-1} = x$ holds iff g and x commute, which in itself is a symmetry statement.

Exercise 10.8. Prove the second statement, that is for a non-trivial group G the action of G on itself by conjugation is also not transitive.
Hint! Show that the identity and a non-identity element can never be in the same orbit.

Solution

(6) If $\rho : G \to \mathcal{S}_X$ is permutation representation of a group action, then G also acts on the power set of X, $\mathcal{P}(X)$ by sending a subset of X to its image under the map $\rho^\bullet$, i.e. $g \bullet_\rho A := \rho(A)$. In symbols

$$g \bullet A = g \bullet_{\mathcal{P}} A = \{g \bullet a : a \in A\}.$$

(7) If we have that $g \bullet A = A$ for every $g \in G$ (i.e. if $G_A = G$), then we get an action $\bullet|_A \colon G \circlearrowright A$ by the same formula as was defined on X. That is for all $x \in A$

$$g \bullet|_A x = g \bullet x.$$

For instance, for the action of G on the set of subsets of G by conjugation, G stabilizes the set of subgroups of G. This is because if $H \leqslant G$, then $gHg^{-1} \leqslant G$ is as well. Therefore

$$G \circlearrowright \{\text{Subgroups of } G\}$$

via $g \bullet H = gHg^{-1}$ is a group action. By definition, the stabilizer of a subgroup H, G_H, is the normalizer of H in G, $\mathrm{N}_G(H)$ which was defined in Exercise 8.40. We also have that H is fixed by every $g \in G$, i.e.

$$H \in \{\text{Subgroups of } G\}^G,$$

iff $H \trianglelefteq G$.

(8) If we have a group action $\bullet \colon G \circlearrowright X$, and a group homomorphism $\varphi \colon \widetilde{G} \to G$, then we can define an action $\star \colon \widetilde{G} \circlearrowright X$ by defining

$$h \star x := \varphi(h) \bullet x.$$

To see this is a group action note

(a) For all $x \in X$

$$\begin{aligned} e_H \star x &= \varphi(e_{\widetilde{G}}) \bullet x && \text{(Def'n of } \star\text{)} \\ &= e_G \bullet x && (\varphi \text{ is a hom of gps} \Rightarrow \text{id of } \widetilde{G} \mapsto \text{id of } G.) \\ &= x. && (\bullet \text{ is a gp action, } \Rightarrow \text{id acts as id}) \end{aligned}$$

(b) Let $h_1, h_2 \in \widetilde{G}$ and $x \in X$ be arbitrary. Then

$$\begin{aligned} (h_1h_2) \star x &= (\varphi(h_1h_2)) \bullet x && \text{(Def'n of } \star\text{)} \\ &= (\varphi(h_1)\varphi(h_2)) \bullet x && (\varphi \text{ a gp hom}) \\ &= \varphi(h_1) \bullet (\varphi(h_2) \bullet x) && (\bullet \text{ is a gp action}) \\ &= h_1 \star (\varphi(h_2) \bullet x) && \text{(Def'n of } \star\text{)} \\ &= h_1 \star (h_2 \star x). && \text{(Def'n of } \star\text{)} \end{aligned}$$

Therefore $\star$ satisfies the two parts of what is needed to be an action, and so it **is** an action.

Actually we could saved a bit of time here since the induced representation of $\star$ is easily seen to be the composition

$$\widetilde{G} \xrightarrow{\varphi} G \xrightarrow{\rho^{\bullet}} \mathcal{S}_X$$

which as the composition of two homomorphisms is a homomorphism. Thus the induced representation of $\star$ is a homomorphism and so $\star$ itself must be an action.

For example, if $H \leqslant G$, then the inclusion map $H \to G$ which sends every element of h to itself, is a homomorphism of groups, and so whenever we have a group action G on a set X we also have a restricted group action of $H \leqslant G$.

Another special case of this comes about when ever consider the Cayley action of $\bullet_R \colon G \circlearrowright G$ (in this paragraph, unlike the previous, we're simply assuming that $\widetilde{G}$ is some group and φ is some homomorphism, we are **not** assuming that $\widetilde{G} \leqslant G$ in this paragraph). In more detail we have an action $\widetilde{G} \circlearrowright G$

$$h \star g := \varphi(h)g$$

where the multiplication on the RHS is defined to be taking place in the group G. In this action, the stabilizer of e_G is

$$\begin{aligned} \widetilde{G}_{e_G} &= \{h \in \widetilde{G} : h \star e_G = e_G\} \\ &= \{h \in \widetilde{G} : \varphi(h)e_G = e_G\} \\ &= \{h \in \widetilde{G} : \varphi(h) = e_G\} \\ &= \operatorname{Ker}(\varphi), \end{aligned} \tag{3}$$

the kernel of the homomorphism φ. The orbit of e_G is

$$\begin{aligned} \widetilde{G} \star e_G &= \{g \in G : \exists h \in \widetilde{G} \text{ s.t. } h \star e_G = g\} \\ &= \{g \in G : \exists h \in \widetilde{G} \text{ s.t. } \varphi(h) = g\} \\ &= \operatorname{Im}(\varphi), \end{aligned} \tag{4}$$

the image of $\varphi \colon \widetilde{G} \to G$.

Exercise 10.9. In the general situtation above, prove that in general

$$\widetilde{G}_x = \varphi^{-1}(G_x) \quad \text{and} \quad \widetilde{G} \star x = \varphi(\widetilde{G}) \bullet x.$$

SOLUTION

(9) Let X be any set. Then any cyclic group of order n, $\langle t \rangle$, acts on the set of n-tuples of X

$$X^n = \{(x_1, \ldots, x_n) : x_1, \ldots, x_n \in X\}$$

by cyclically permuting the $x_i's$, i.e.

$$t \bullet (x_1, \ldots, x_n) = (x_2, x_3, \ldots, x_n, x_1)$$

(10) Let X be any set. $\mathcal{S}_n$ acts on the set of n-tuples of X

$$\{(x_1, \ldots, x_n) : x_1, \ldots, x_n \in X\}$$

by permuting the $x_i's$, i.e.

$$\sigma \bullet (x_1, \ldots, x_n) = (x_{\sigma^{-1}(1)}, \ldots, x_{\sigma^{-1}(n)})$$

To see that this defines an action note that the identity obviously acts as the identity. Next, let $\sigma, \tau \in \mathcal{S}_n$ and $(x_1, \ldots, x_n) \in X^n$ be arbitrary. Let's define y_i as $x_{\tau^{-1}(i)}$.

$$\begin{aligned}\sigma \bullet \Big(\tau \bullet (x_1, \ldots, x_n)\Big) &= \sigma \bullet (x_{\tau^{-1}(1)}, \ldots, x_{\tau^{-1}(n)}) \\ &= \sigma \bullet (x_{\tau^{-1}(1)}, \ldots, x_{\tau^{-1}(n)}) \\ &= \sigma \bullet (y_1, \ldots, y_n) \\ &= (y_{\sigma^{-1}(1)}, \ldots, y_{\sigma^{-1}(n)}).\end{aligned}$$

Now let $1 \leq i \leq n$ (i.e. $i \in \overline{n}$) be an integer and let $j = \sigma^{-1} i$. Then

$$y_{\sigma^{-1}(i)} = y_j = x_{\sigma^{-1}(j)} = x_{\sigma^{-1}(\sigma^{-1}(i))}.$$

Hence we have that

$$\begin{aligned}\sigma \bullet (\tau \bullet (x_1, \ldots, x_n)) &= (x_{\tau^{-1}(\sigma^{-1}(1))}, \ldots, x_{\tau^{-1}(\sigma^{-1}(n))}) \\ &= (x_{(\sigma\tau)^{-1}(1)}, \ldots, x_{(\sigma\tau)^{-1}(n)}) \\ &= (\sigma\tau) \bullet (x_1, \ldots, x_n).\end{aligned}$$

EXERCISE 10.10. Suppose that $G \curvearrowright X$ and let n be a positive integer. Prove that the coordinate action of $G \curvearrowright X^n$ and the action of $\mathcal{S}_n \curvearrowright X^n$ from the previous example commute.

SOLUTION

(11) Every group of matrices of degree n and coefficients in a ring R acts on the set of n-vectors,

$$R^n = \left\{ \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} : r_i \in R \right\}.$$

For example, in the case that $n = 2$, the action is via

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \bullet \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}.$$

The action is faithful since only the identity matrix fixes all vectors. Note that every matrix fixes the zero vector.

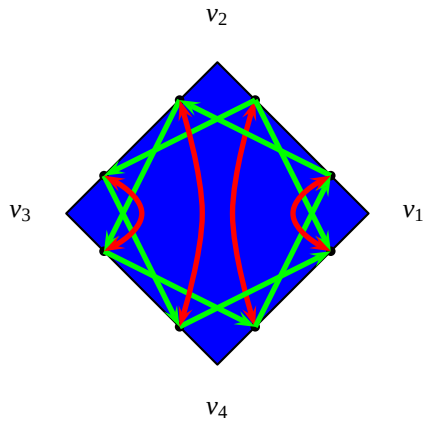
(12) For any action $G \curvearrowright X$, if G stabilizes a subset $A \subseteq X$, meaning for all $a \in A$ and $g \in G$, $g \bullet a \in A$, then G also acts on A. Indeed both axioms of a group action have "for all" quantifiers in them. Note that since the ρ_g's are bijections, we also have that G preserves $X \setminus A$ whenever G preserves X, so G also acts on the compliment $X \setminus A$.

Here are some sub-examples of this:

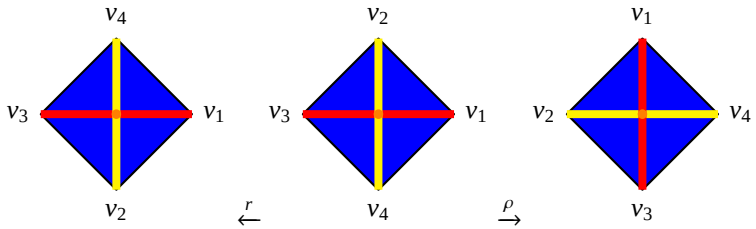
(a) If G preserves two sets A_1 and A_2, then G also preserves their intersection, union, compliments, and even their symmetric difference. For example, if $a \in A_1 \cap A_2$, then $a \in A_1$ so for all $g \in G$, $g \bullet a \in A_1$, and similarly $g \bullet a \in A_2$, hence $g \bullet a \in A_1 \cap A_2$ whenever a is. A similar proof applies to $A_1 \cup A_2$, $A_1 \setminus A_2$, $A_1 \Delta A_2$, etc.

(b) $A \subseteq X$ is preserved by G iff A is a union of orbits of $G \circlearrowright X$. Indeed, for any $y \in G \bullet x$, $y = \overline{g} \bullet x$, we have that $g \bullet y = g\overline{g} \bullet x$ which is still in the orbit of x. The G preserves a union of its orbits. On the other hand, if x is in some preserved subset of X, then for all $g \in G$, $g \bullet x \in A$. But the set of all $g \bullet x$'s is, by its very definition, the orbit of x. Hence for any $x \in A$, $G \bullet x \subseteq A$.

(c) We have $\mathcal{D}_4 \circlearrowright \blacklozenge^2$, but $\mathcal{D}_4$ preserves the four vertices of the square as well as the four centers of edges. Moreover it preserves the center of the square. All other preserved subsets can be seen to have 8 elements in it. For instance, if we take a point halfway between the midpoint of an edge and a vertex.



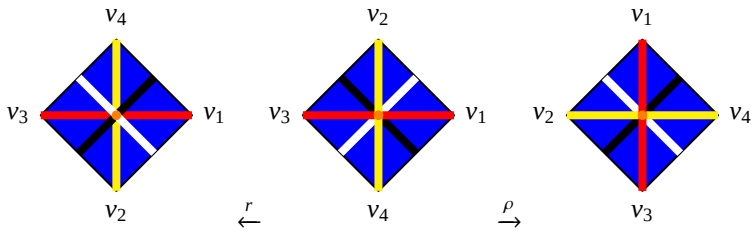
(d) Since G also acts on the subsets of X, we can apply the same theory to $\mathcal{P}(X)$. For instance the set of edges of a square are preserved by $\mathcal{D}_4$. As well as the pair of diagonals of the square.



Therefore, we have the permutation representation (using the obvious dictionary)

$$\rho \mapsto \begin{pmatrix} x\text{ axis} & y\text{ axis} \\ y\text{ axis} & x\text{ axis} \end{pmatrix} = (\mathbf{12})$$
$$r \mapsto \begin{pmatrix} x\text{ axis} & y\text{ axis} \\ x\text{ axis} & y\text{ axis} \end{pmatrix} = (\mathbf{1})(\mathbf{2}).$$

As such the reflection r is the kernel of this action, but so is ρ^2. In fact, the kernel of this action is $\left\langle \rho^2, r \right\rangle$. If we throw in the diagonals which connect the edges as well then things look as follows:



This time we get the permutation representation (again, we're using the obvious dictionary)

$$\rho \mapsto \begin{pmatrix} x\,\text{axis} & y\,\text{axis} & x+y\,\text{axis} & x-y\,\text{axis} \\ y\,\text{axis} & x\,\text{axis} & x-y\,\text{axis} & x+y\,\text{axis} \end{pmatrix} = (\mathbf{12})(\mathbf{34})$$
$$r \mapsto \begin{pmatrix} x\,\text{axis} & y\,\text{axis} & x+y\,\text{axis} & x-y\,\text{axis} \\ x\,\text{axis} & y\,\text{axis} & x-y\,\text{axis} & x+y\,\text{axis} \end{pmatrix} = (\mathbf{34}).$$

This time only the 180 degree rotation, ρ^2, is a non-trivial element in the kernel, so the kernel is $\langle \rho^2 \rangle$. This action is, however, now no longer transitive since the axis connecting the edges and the vertices cannot be moved to each other via a symmetry of a square.

(e) In three dimensions, the symmetries of a cube preserve the three coordinate axis of a cube. In this way we get an action of $S\left(\blacklozenge^3\right)$ on a set with three elements. This action is not faithful. For instance, the map

$$(x, y, z) \mapsto (-x, -y, z)$$

is a symmetry of a cube which corresponds to the rotation $\varrho_{\pi,\mathbf{k}}$. This clearly fixes all three coordinate axis, and is therefore in the kernel of this action. It should be noted that (using the dictionary of labeling the x axis as 1, y axis as 2, and z axis as 3.)

$$\varrho_{\pi/2,\mathbf{k}} \mapsto (\mathbf{12})$$
$$\varrho_{\pi/2,\mathbf{i}} \mapsto (\mathbf{23})$$
$$\varrho_{\pi/2,\mathbf{j}} \mapsto (\mathbf{13}).$$

This permutation representation is clearly surjective (onto $\mathcal{S}_3$), as can be seen by either noting that we have already hit a generating set of $\mathcal{S}_3$, or as seen more directly by writing out the image of more symmetries. For instance,

$$\varrho_{\frac{2\pi}{3},\mathbf{i}+\mathbf{j}+\mathbf{k}} = \varrho_{\pi/2,\mathbf{i}} \circ \varrho_{\pi/2,\mathbf{j}} \mapsto (\mathbf{23})(\mathbf{12}) = (\mathbf{123}).$$

The kernel can either seen directly, or by using **KIT** (to be proven below) that the kernel of this action consists of (in addition to the identity) the three 180° degree rotations about the faces.

(f) Now let's consider the action of $S\left(\blacklozenge^3\right)$ on the set of 4 diagonals which connect the vertices of the cube. It can be seen with a bit of geometry the only symmetry which fixes these four axis is the antipolar map, $(x, y, z) \mapsto (-x, -y, -z)$, but this map is not orientation preserving, so we do not need to worry about. Thus the action of $S\left(\blacklozenge^3\right)$ on these 4 axis is faithful. Therefore we have an injective homomorphism

$$S\left(\bigstar^3\right) \to \mathcal{S}_4$$

which is injective. However both of these groups have order 24, so this homomorphism is actually surjective as well! Thus we have proven:

THEOREM 10.14. *The groups $S\left(\bigstar^3\right)$ and $\mathcal{S}_4$ are isomorphic.*

Let's define the dictionary

$$\begin{aligned} 1 &:= \text{line connecting } (\ 1,\ \ 1,\ \ 1) \text{ to } (-1,-1,-1),\\ 2 &:= \text{line connecting } (-1,\ \ 1,\ \ 1) \text{ to } (\ 1,-1,-1),\\ 3 &:= \text{line connecting } (\ 1,-1,\ \ 1) \text{ to } (-1,\ \ 1,-1),\\ 4 &:= \text{line connecting } (\ 1,\ \ 1,-1) \text{ to } (-1,-1,\ \ 1). \end{aligned}$$

Since the map

$$\varrho_{\pi/2,\mathbf{i}} \colon (x,y,z) \mapsto (x,-z,y)$$

we have that

$$\begin{aligned} \varrho_{\pi/2,\mathbf{i}}(1,1,1) &= (1,-1,1) & \varrho_{\pi/2,\mathbf{i}}(-1,1,1) &= (-1,-1,1)\\ \varrho_{\pi/2,\mathbf{i}}(1,-1,1) &= (1,-1,-1) & \varrho_{\pi/2,\mathbf{i}}(1,1,-1) &= (1,1,1). \end{aligned}$$

Thus in terms of the permutation representation, we have that

$$\varrho_{\pi/2,\mathbf{i}} \mapsto (\mathbf{1\ 3\ 2\ 4}).$$

Similarly we also have that

$$\varrho_{\pi/2,\mathbf{j}} \mapsto (\mathbf{1\ 2\ 3\ 4}) \quad \text{and} \quad \varrho_{\pi/2,\mathbf{k}} \mapsto (\mathbf{1\ 2\ 4\ 3}).$$

REMARK 10.15. The presence of the inverse in (10) might seem strange on first glance, but this is needed and looks much more natural when writing a specific example. For instance

$$(1\ 2\ 3) \bullet (x_1,x_2,x_3) = (x_3,x_1,x_2)$$

says that the first term goes to the second term (as it should) but the first *location* has its new input given by what was the input of the third location. It has to do with the fact that we compose functions from right to left instead of from left to right.

EXERCISE 10.11. Suppose we have two actions, $\bullet_1 \colon G \circlearrowright X$ and $\bullet_2 \colon H \circlearrowright X$ which commute in the sense that

$$g \bullet_1 (h \bullet_2 x) = h \bullet_2 (g \bullet_1 x)$$

for all $g \in G$ and $h \in H$.

(1) Prove that $(g,h) \bullet x := g \bullet_1 (h \bullet_2 x)$ defines an action $\bullet \colon G \times H \circlearrowright X$.

(2) Prove that $g \bullet (H \bullet x) := H(g \bullet x)$ gives a well defined action of G on the set of orbits of X under the action of H.

SOLUTION

Exercise 10.12. Suppose that $\bullet\colon G \circlearrowright X$.

Solution

(1) Show that G also acts on $\mathcal{P}_2(X)$ via

$$g \bullet' \{x_1, x_2\} := \{g \bullet x_1, g \bullet x_2\}.$$

Generalize this by considering the set

$$\mathcal{P}_n(X) \subseteq \mathcal{P}(X).$$

(2) Instead, define $G\colon X \times X$, also denote $\bullet$, which is similar to the last problem:

$$g \bullet (x_1, x_2) := (g \bullet x_1, g \bullet x_2).$$

Prove that $\bullet$ is an action. Generalize this by considering the set of n-tuples of elements in X,

$$X^n = X \times \cdots \times X.$$

Show that G also acts on the sets (when X has more than n elements)

$$X^{*n}, \qquad \mathcal{P}(X), \quad \text{and} \quad \mathcal{P}_n(X).$$

Definition 10.16. Recall an action, $G \circlearrowright X$, with only one orbit is said to be *transitive* and in this case G is said to act *transitively* on X. More generally, if the induced action of G on X^{*n} [III].

is transitive, the action is said to be *n-transitive*

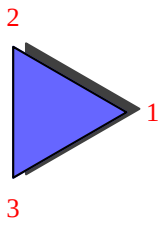
Remark 10.17. For any set X, $X^{*1} = X$. Thus, requiring an action be 1 transitive is the same as requiring that an action be transitive.

In less fancy language, an action is transitive if any two points of X can be lined up by an element of G. That is for any x and y in X there exists a $g \in G$ so that $g \bullet x = y$. The action is 2-transitive if any pair of pairs of points can be lined up, 3-transitive if any pair of triples can be lined up etc.

Example 10.18. The action of $\mathcal{D}_4 \circlearrowright$  gives us a transitive action on vertices, but this action is not two transitive since there is no symmetry which takes a pair of opposite vertices to a pair of adjacent vertices.

On the other hand the action of $\mathcal{D}_3 \circlearrowright$  gives us a 1, 2, and 3-transitive action on the set of vertices of the triangle.

[III] Recall, this is just the set of ordered n-tuples in X with no two equal. So if $X = \overline{3} = \{1, 2, 3\}$, then $(1, 2) \in X^{*2}$ but $(1, 1)$ is not.



Subgroups of $\mathcal{S}_n$ which act n-transitively $\overline{n}$ are called *n-transitive subgroups* of $\mathcal{S}_n$. Notice that Cayley's theorem states that every finite group of order n is a transitive subgroup of $\mathcal{S}_n$.

THEOREM 10.19. *For any action of G on X*

(1) *The relation*

$$x \sim y \Leftrightarrow \exists g \in G \text{ s.t. } g \bullet x = y$$

is an equivalence relation. The equivalence class of x is the orbit of x, $G \bullet x$.

(2) *The action of G restricts to a transitive action on every orbit.*

PROOF.

(1) We show that the relation is reflexive symmetric and transitive.

- For all $x \in X$, $e \bullet x = x$. Thus the relation is reflexive.
- $g \bullet x = y$ holds iff $g^{-1} \bullet y = x$. Hence the relation is symmetric.
- $g \bullet x = y$ and $h \bullet y = z$ implies that $hg \bullet x = z$. Hence the relation is transitive.

(2) Holds by definition. ∎

This should be thought of as saying that every action is the disjoint union of transitive actions.

EXAMPLE 10.20. $\mathcal{S}_n$ is an n-transitive subgroup of itself. Indeed, for any two n-tuples

$$(a_1, \ldots, a_n) \quad \text{and} \quad (b_1, \ldots, b_n)$$

of distinct points define

$$\sigma := \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix} \quad \text{and} \quad \tau := \begin{pmatrix} 1 & 2 & \cdots & n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix},$$ IV

then

$$\sigma \bullet (1, \ldots, n) = (a_1, \ldots, a_n) \Rightarrow \sigma^{-1} \bullet (a_1, \ldots, a_n) = (1, \ldots, n)$$

and

$$\tau \bullet (1, \ldots, n) = (b_1, \ldots, b_n) \Rightarrow \left(\tau\sigma^{-1}\right) \bullet (a_1, \ldots, a_n) = (b_1, \ldots, b_n).$$

This concludes the proof that $\mathcal{S}_n \circlearrowright \overline{n}$ is n-transitive since $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ were arbitrary.

Meanwhile, note that while $\mathcal{A}_3$ does act transitively on $\overline{3}$, it does not act 2-transitively. Indeed, $(1\ 2\ 3)$ is the only even permutation in $\mathcal{S}_3$ which can take 1 to 2. This permutation takes 2 to 3. Therefore, it's impossible to find an even permutation on three letters which takes $(1, 2)$ to $(2, 1)$.

EXERCISE 10.13. Prove that if $G \circlearrowright X$ is transitive, then the only (non-empty) invariant subset of X is X itself.

SOLUTION

The following exercise shows that requiring an action be two–transitive is not the same as requiring the action on pairs be transitive (it is requiring that the action on ordered pairs be transitive.)

EXERCISE 10.14. Consider $\mathbb{Z}_3 \circlearrowright \{x, y, z\}$ via the 3–cycle $(x\ y\ z)$.[V] Prove that

SOLUTION

(1) This action is not 2–transitive.

(2) This action is transitive on pairs of elements:

$$\mathcal{P}_2(\{x, y, z\}) = \{\{x, y\}, \{x, z\}, \{y, z\}\}.$$

EXERCISE 10.15. Prove that the action of $\mathcal{A}_4 \circlearrowright \overline{4}$ is 2-transitive, but it is not 3-transitive.

SOLUTION

EXERCISE 10.16.

SOLUTION

(1) Suppose that $G \leqslant \mathcal{S}_n$ is an n-transitive subgroup. Prove that $G = \mathcal{S}_n$.

(2) Suppose that $\#X = n$. Suppose that $G \circlearrowright X$ is $n-1$-transitive. Prove that the action is also n-transitive.
Hint! Any injection $\iota\colon \overline{n-1} \to \overline{n}$ determines a unique bijection $\widetilde{\iota}\colon \overline{n} \to \overline{n}$ by declaring that $\widetilde{\iota}(n)$ is equal to the unique point of $\overline{n}$ not in the image of $\overline{n}$.

(3) Suppose that $G \leqslant \mathcal{S}_n$ is an $n-1$-transitive subgroup. Prove that $G = \mathcal{S}_n$.

[IV] These dudes really are bijections because we are assuming that $a_i = a_j \Rightarrow i = j$ and the same for the $b's$.
[V] Meaning $[1]_3 \bullet x = y$, $[1]_3 \bullet y = z$, and $[1]_3 \bullet z = x$.

Exercise 10.17. Prove that ($n > 1$)

(1) $\mathcal{A}_n$ is an $n-2$ transitive subgroup of $\mathcal{S}_n$,

(2) $\mathcal{A}_n$ is not an $n-1$ transitive subgroup of $\mathcal{S}_n$.

SOLUTION

Hint! There are two possibilities here

(1) For $n > 2$ it's easy to see that the action of $\mathcal{A}_n \circlearrowright \overline{n}$ is transitive. To prove the more general result note that in $\mathcal{A}_n$ the stabilizer of the tuple $(n, n-1, n-2, \ldots, m+1)$ is $\mathcal{A}_m$.

(2) For an arbitrary $n-2$-tuple of distinct points

$$(a_1, \ldots, a_{n-2})$$

use a permutation $\sigma \in \mathcal{S}_n$ to move $(1, \ldots, n-2)$ onto it. Now the only thing that could go wrong occurs if σ happened to be an odd permutation. In that dire and awful situation, consider instead the permutation $\sigma \circ (n-1\ n)$. You still need to check that this new permutation is an even permutation and takes

$$1 \mapsto a_1 \ldots n-2 \mapsto a_{n-2}.$$

(My wording here sucks, there are much more concise ways of writing this up.) To prove that $\mathcal{A}_n$ is not an $n-1$–transitive subgroup of $\mathcal{S}_n$ use the previous exercise.

Exercise 10.18. Suppose that $G \circlearrowright X$ is a group action and $x \in X$. Define the orbit map

$$\begin{aligned} \alpha_x &: G \to X \\ &: g \mapsto g \bullet x. \end{aligned}$$

SOLUTION

Prove that

(1) $\mathrm{Im}(\alpha_x) = G \bullet x$. In particular, $G \circlearrowright X$ is transitive iff α_x is surjective;

(2) α_x is injective iff the stabilizer of x is trivial;

(3) α_x is a bijection iff $G \circlearrowright X$ is free and transitive.

Exercise 10.19. Suppose that $G \circlearrowright X$ is a transitive action of a finite group G on a set X. Prove the following are equivalent

SOLUTION

(1) Some $x \in X$ has a trivial stabilizer. That is for some $x \in X$, $G_x = \{e_G\}$,

(2) Every $x \in X$ has a trivial stabilizer. That is for all $x \in X$, $G_x = \{e_G\}$;

(3) The action $G \circlearrowright X$ is free;

(4) $|G| = \#X$.

Hint! There is quite a bit of overlap between this problem and previous one.

DEFINITION 10.21. The set of orbits of an action of G on X is denoted X/G and is called the *quotient space*.

For a fixed generating set S of G, a useful way of picturing an action is through an action diagram. It is just the union of the digraphs of the various generators colored differently for the different generators. That is the digraph has vertices X and we connect two vertices x_1 and x_2 by an edge "colored" $s \in S$ if we have $s \bullet x_1 = x_2$.

EXERCISE 10.20. Show that an action of $G \circlearrowright X$ is transitive iff its action diagram is connected for some generating set $S \subseteq G$ iff its action diagram is connected for any generating set $S \subseteq G$.

SOLUTION

EXAMPLE 10.22. Consider the action of $\mathcal{D}_4$ on itself by conjugation. Then we have the diagram {Draw}

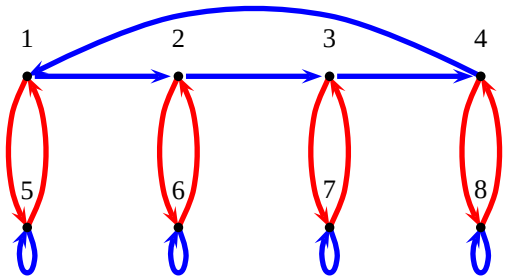
EXAMPLE 10.23. The group

$$\left\langle (1\,\overset{x}{2}\,3\,4), (1\,5)(2\,6)\overset{y}{(}3\,7)(4\,8) \right\rangle,$$

which we denote as G, is a transitive subgroup of $\mathcal{S}_8$ since

$x(1) = 2$ $x(2) = 3$ $x(3) = 4$ $y(4) = 8$
$y(1) = 5$ $y(2) = 6$ $y(3) = 7$

shows that every letter is in the same orbit as 1. This can be best seen by the diagram



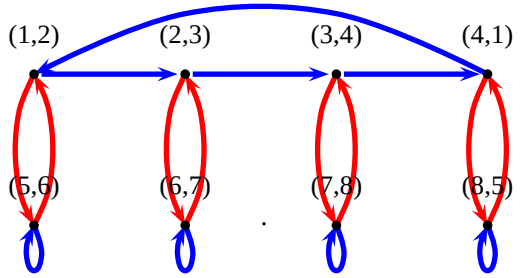
However, the action is not 2-transitive. More specifically some playing {perhaps I draw this out?? } shows that there are five total orbits which are represented by

$$\overline{8}^{*2}/_G = \{G \bullet (1,2), G \bullet (1,3), G \bullet (1,4), G \bullet (1,5)\}$$

of the action of G on $\overline{8}^{*2}$. (The first three orbits have diagrams which look like the one above and have 8 elements each while the last orbit, the orbit of $(1,5)$ has 32 elements in it ... this matches the fact there are $3 \cdot 8 + 32 = 56$ elements in $\overline{8}^{*2}$). For instance the the orbit of $(1,2)$,

$$G \bullet (1,2) = \{(1,2),(2,3),(3,4),(4,1),(5,6),(6,7),(7,8),(8,1)\}$$

as depicted as



(We note that we have the entire orbit because every point has two arrows coming from it. Thus the diagram is invariant under the action of the generators and hence invariant under the action of the entire group). We can see directly that $(2,1)$ is not in the orbit of $(1,2)$. This is simply because we just drew out the entire orbit of $(1,2)$ and $(2,1)$ didn't end up showing up! We note now that it is **not** a coincidence that these two diagrams look exactly alike (up to how we name the vertices). We will see that it's not a coincidence when we prove the orbit stabilizer theorem. That theorem says the fact that these two diagrams look exactly the same is due to the fact that the stabilizer of the point 1 is exactly the same as the pair of points, $(1,2)$. Indeed, with a bit more effort, we will see that these stabilizers each consist only of powers of x. That is, written in symbols:

$$G_1 = \langle x \rangle = G_{(1,2)}.$$

Next, let's note that the action of $G \circlearrowright (G \bullet (1,5))$ is free. This is due to the fact that the original diagram above displays a symmetry via $x \bullet$. In more detail, from this symmetry we have that the stabilizer of 1 is the same as the stabilizer of $x \bullet 1 = 2$ which is the stabilizer of 3 and 4. On the other hand, the stabilizer of 5 is the same as that of 6, which is the same as that of 7, which is the same as that of 8. Hence if $g \in G$ fixes the ordered pair $(1,5)$, then (as always) it must fix 1 and 5, which, since it fixes 1 it must fix also 2, 3, and 4; and since it fixes 5 it must also fix 6, 7, and 8. Therefore, if a $g \in G$ fixes both letters, 1 and 5, then it must fix every letter. That is it must be the identity. Therefore, the stabilizer of $(1,5)$ is trivial. Since the action on an orbit is by definition transitive, the stabilizer of every point in the orbit of $(1,5)$ must also be trivial. Therefore only the identity can fix a point in the orbit of $(1,5)$, which is simply another way of saying that the action of G on the orbit of $(1,5)$ is free.

Now let's examine how many points are in the orbit of $(1,5)$. Let's define

$$X_1 := \{(i,j) : i \in \{1,2,3,4\} \text{ and } j \in \{5,6,7,8\}\}$$
$$X_2 := \{(i,j) : j \in \{1,2,3,4\} \text{ and } i \in \{5,6,7,8\}\}.$$

Of course $X_1 \cap X_2 = \varnothing$. I claim that $G \bullet (1,5) = X_1 \cup X_2$. For ease of notation, denote z as the conjugate of x by y, that is $z := yxy^{-1}$. Now the effect of x and z is easy to see. For example

$$x^3 z^2 \bullet (1,5) = (4,7).$$

Indeed, if we want to get to a point $(i,j) \in X_1$ from $(1,5)$, simply hit $(1,5)$ by $x^{i-1} z^{j-1}$. This is because $x = (1\ 2\ 3\ 4)$ pushes the numbers 1 through 4 forward by 1 while fixing the numbers 5 through 8 while $z = (5\ 6\ 7\ 8)$ does the oppositie: it pushes the numbers 5 through 8 forward by 1 while fixing the numbers 1 through 4. In particular,

$\langle x, z \rangle \leqslant G$ stabilizers the sets X_1 and X_2 and acts transitively (and freely) on each individually. How do we get from points in X_1 to points in X_2 (and back?). Simply hit by y first, which takes $(1, 5)$ to $(5, 1)$, and then proceed as before. For instance how do we get to $(7, 2)$? Well,

$$(z^2xy) \bullet (1, 5) = (7, 2)$$

does the trick.

How many points are in $X_1 \cup X_2$? Well, for every $1 \leq i \leq 4$, there are four points $(i, 5)$, $(i, 6)$, $(i, 7)$, and $(i, 8)$ in X_1. Thus there are 16 points in X_1, and so there are also 16 points in X_2. Hence there are 32 points in the orbit of $(1, 5)$. Therefore G acts freely and transitively on a set with 32 elements. This implies, among other things, that $|G| = 32$.

EXERCISE 10.21. Concerning Example 10.23, which of the five orbits is $(2, 1)$ in? (It's not in the orbit of $(1, 2)$).

SOLUTION

THEOREM 10.24. *Suppose G acts on X, then*

(1) *The action is n-transitive iff*

- *The action is $n-1$-transitive* [VI]
- *For some $\mathbf{x} = (x_1, \ldots, x_n) \in X^{*n}$ we have $G_{\mathbf{x}}$ acts transitively on $X^{*n} \setminus \{\mathbf{x}\}$.*

(2) *For any subset of $A \subseteq X$ we have $G_A \leqslant G$.*

PROOF.

(1) For notational ease, we do this in the special case that $n = 2$ and leave the more general – but entirely analogous – case to the reader.

$\Rightarrow$ Suppose that $G \circlearrowright X$ is 2–transitive and let x and y be distinct points in X. Since the action is 2–transitive, there exists a $g \in G$ with $g \bullet (x, y) = (y, x)$. But this means that $g \bullet x = y$ and the action is transitive. We now show that the action of G_x is transitive on $X \setminus \{x\}$. Indeed, if we have distinct $y, z \in X$–neither equal to x, then since G's action on X is 2-transitive there exists a $g \in G$ with

$$g \bullet (x, y) = (x, z).$$

This is the case since both (x, y) and (x, z) are pairs of distinct points.

$\Leftarrow$ Let $x \neq y$ and $z \neq w$. We can assume that G_x acts transitively on $X \setminus \{x\}$ (why?). Since the action of G on X is transitive, there exists a $b \in G$ with $b \bullet z = x$. Moreover, since the action of G_x is transitive on $X \setminus \{x\}$ there exists a $a \in G$ with $a \bullet x = x$ and $a \bullet (b \bullet z) = y$. Then $g = ab$ satisfies $g \bullet (z, w) = (x, y)$ and, by definition, this means that the action of G on X is 2-transitive. (This proof could be titled "Line it up and fire away"!)

(2) We apply the two step subgroup test.

TS1 Let $g, h \in G_A$. Then

$$(gh) \bullet A = g \bullet (h \bullet A) = g \bullet A = A.$$

Hence G_A is closed under the operation.

[VI] We need to assume that X has more than n elements!

TS2 Let $g \in G_A$. Then

$$\begin{aligned} A &= e \bullet A \\ &= g^{-1}g \bullet A \\ &= g^{-1} \bullet (g \bullet A) \\ &= g^{-1} \bullet A. \end{aligned}$$

Hence $g^{-1} \in G_A$ as well. ■

EXERCISE 10.22. Suppose that G is generated by a set S and $G \circlearrowright X$. For $x \in X$, denote the generators (in S) which fix x as S_x, in symbols:

$$S_x := \{s \in S : s \bullet x = x\}.$$

Show that $\langle S_x \rangle \preccurlyeq G_x$ but equality needn't hold. (In fact equality rarely holds!)

SOLUTION

MAPLE EXAMPLE 10.25. As usual, we let ✶^{3} be the icosahedron (at times we let it denote its dual, the dodecahedron) If we pick a dictionary that labels the 6 lines through the origin which connect the 6 pairs of opposite vertices of the icosahedron, we get an action of $S\left(✶^3\right) \circlearrowright \overline{6}$. It is easy to see that $S\left(✶^3\right)$ is generated by any two rotations above adjacent vertices. In terms of a dictionary these two permutations are written in cycle form as

$$(2\,3\,4\,5\,6) \quad \text{and} \quad (1\,3\,5\,4\,6).$$

Cheating, and using `MAPLE` (and hence using their notation for a permutation written in cycle notation) we see that this consists of the permutations

[[ ]]

[[1, 2, 3, 5, 6]], [[1, 2, 4, 3, 5]], [[1, 2, 5, 6, 4]], [[1, 2, 6, 4, 3]],
[[1, 3, 2, 5, 4]], [[1, 3, 4, 6, 2]], [[1, 3, 5, 4, 6]], [[1, 3, 6, 2, 5]],
[[1, 4, 2, 3, 6]], [[1, 4, 3, 6, 5]], [[1, 4, 5, 2, 3]], [[1, 4, 6, 5, 2]],
[[1, 5, 2, 6, 3]], [[1, 5, 3, 4, 2]], [[1, 5, 4, 2, 6]], [[1, 5, 6, 3, 4]],
[[1, 6, 2, 4, 5]], [[1, 6, 3, 2, 4]], [[1, 6, 4, 5, 3]], [[1, 6, 5, 3, 2]],
[[2, 3, 4, 5, 6]], [[2, 4, 6, 3, 5]], [[2, 5, 3, 6, 4]], [[2, 6, 5, 4, 3]],
[[1, 2], [3, 6]], [[1, 2], [4, 5]], [[1, 3], [2, 4]], [[1, 3], [5, 6]], [[1, 4], [2, 6]],
[[1, 4], [3, 5]], [[1, 5], [2, 3]], [[1, 5], [4, 6]], [[1, 6], [2, 5]], [[1, 6], [3, 4]],
[[2, 3], [4, 6]], [[2, 4], [5, 6]], [[2, 5], [3, 4]], [[2, 6], [3, 5]], [[3, 6], [4, 5]],
[[1, 2, 3], [4, 6, 5]], [[1, 2, 4], [3, 6, 5]], [[1, 2, 5], [3, 4, 6]], [[1, 2, 6], [3, 4, 5]],
[[1, 3, 2], [4, 5, 6]], [[1, 3, 4], [2, 6, 5]], [[1, 3, 5], [2, 6, 4]], [[1, 3, 6], [2, 4, 5]],
[[1, 4, 2], [3, 5, 6]], [[1, 4, 3], [2, 5, 6]], [[1, 4, 5], [2, 6, 3]], [[1, 4, 6], [2, 5, 3]],
[[1, 5, 2], [3, 6, 4]], [[1, 5, 3], [2, 4, 6]], [[1, 5, 4], [2, 3, 6]], [[1, 5, 6], [2, 4, 3]],
[[1, 6, 2], [3, 5, 4]], [[1, 6, 3], [2, 5, 4]], [[1, 6, 4], [2, 3, 5]], [[1, 6, 5], [2, 3, 4]]

with

- 1 1–cycle, the identity;
- 24 5–cycles, $5 - 1 = 4$ for each of the $12/2 = 6$ pairs of opposite vertices of an icosahedron;
- 15 2 2–cycles, $2 - 1 = 1$ for each of the $30/2 = 15$ pairs of opposite edges of of an icosahedron;
- 20 3 3–cycles, $3 - 1 = 2$ for each of the $20/2 = 10$ pairs of opposite faces of an icosahedron.

REMARK 10.26. In terms of the Rubik's cube, it is possible for us to make moves which fix a face or a cubelet after performing all the moves, but none fix a face or a cubelet along the way. Certain generating sets are better behaved with this sort of thing. For instance, if we take

$$\mathcal{S}_5 = \left\langle \overset{s_1}{\overset{\shortparallel}{(1\,2)}}, \overset{s_2}{\overset{\shortparallel}{(2\,3)}}, \overset{s_3}{\overset{\shortparallel}{(3\,4)}}, \overset{s_4}{\overset{\shortparallel}{(4\,5)}} \right\rangle$$

then we have that $(\mathcal{S}_5)_5 = \langle\, s_1, s_2, s_3 \,\rangle$. However, if instead we take the generating set

$$\left\langle \overset{x}{\overset{\shortparallel}{(1\,2)}}, \overset{y}{\overset{\shortparallel}{(1\,2\,3\,4\,5)}} \right\rangle$$

for $\mathcal{S}_5$, then only one of these generators which fix 5 is x. Clearly $\langle\, x \,\rangle$ can't be the entire stabilizer since it only has order 2, but the stabilizer has order $4! = 24$.

EXAMPLE 10.27.

(1) Any subgroup of $\mathcal{S}_n$ which contains an n–cycle and an $n-1$–cycle must be 2-transitive. This is because the n–cycle alone acts transitively on $\overline{n}$ and the $n-1$–cycle acts transitively on the set of letters it does not omit.

(2) We denote the group

$$\left\langle (1\,2\,3\overset{x}{\overset{\shortparallel}{4}}5\,6\,7), (1\,2)\overset{y}{\overset{\shortparallel}{(}}3\,6) \right\rangle$$

as G. G is clearly a transitive subgroup of $\mathcal{S}_7$ (just x acts transitively on $\overline{7}$). It is also 2-transitive, but to show this is trickier. I'll try to explain my thought process (as opposed to just proving it). We want to show that $G_{\{1\}}$ acts transitively on the set

$$\overline{7} \setminus \{1\} = \{2, 3, 4, 5, 6, 7\}.$$

As it stands, neither generator stabilizes the letter 1, but we should note that both x and y do the same thing to 1. Thus $y^{-1}x = yx$ does nothing to 1. Indeed, a direct

calculation show that

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

Thus we already know that 2,6, and 7 are in the same orbit of the stabilizer and 3,4, and 5 are also in the same orbit. So we know that

$$\{2,6,7\} \subseteq G_{\{1\}} \bullet 2 \quad \text{and} \quad \{3,4,5\} \subseteq G_{\{1\}} \bullet 3.$$

So we just need to find some element of the stabilizer that mixes these two pieces. For instance, we could show 5 is in the orbit of 2, or that 6 is in the orbit of 4 of any number of things (draw a picture of this to get used to dealing with equivalence relations!). Now, if we check out yx^2 we see that this equals a $\underline{7}$–cycle, which fixes squat, so it's of no use. But we are in luck since yx^3 is the $\underline{4}\,\underline{2}$–cycle $(1\,4\,7\,6)(2\,5)$. You should be thinking, this doesn't fix 1 it fixes 3, so I don't care. But alas, if we conjugate this guy by x^{-1} it will fix 2, then if we conjugate it again by x^{-1} it will end as a $\underline{4}\,\underline{2}$–cycle which does indeed fix 1. We know we're done already (without having to do any further conjugations) since we know that four cycle part must contain both elements of the first set $\{2\,6\,7\}$ and the second set $\{3\,4\,5\}$ by the pigeon hole principle. Thus this dude must fix these two halves like we wanted. But, all right, I do this out ...

$$x^{-2}yx^3x^2 = (6\,2\,5\,4)(7\,3) = (2\,5\,4\,6)(3\,7)$$

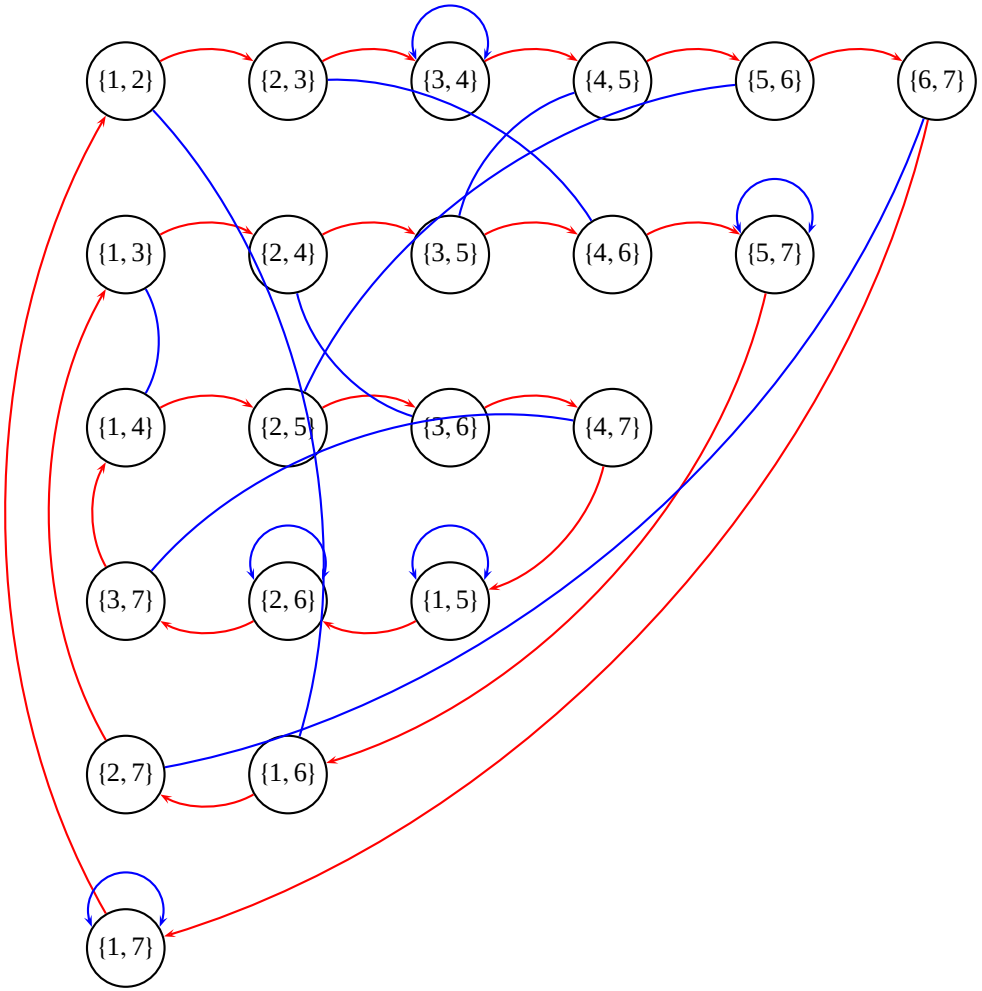
by the $\mathcal{S}_n$ conjugacy trick. Thus we have that 5 is in the orbit of 2 (or if you prefer 2 is in the orbit of 5) and so we are done. It turns out this group is not 3-transitive, but in its current form this problem is too long.

(3) {Sync these examples}

Similar to the previous example, we draw the action diagram for the action

$$G = \langle (1\,2\,3\,4\,5\,6\,7), (2\,6)(3\,4) \rangle$$

on the $\binom{7}{2} = 21$ elements of $\mathcal{P}_2(\overline{7})$. (red arrows correspond to the action of the first generator and blue arrows correspond to the action of the second. Notice that this diagram shows that this action is transitive. Indeed, you simply need to see that the three red loops have elements which can be moved to each other.



(4) The symmetric group (and hence any of its subgroups) acts on the set of polynomials in n–variables by permuting these variables according to

$$\sigma \bullet x_i = x_{\sigma(i)}$$

and extending in the obvious way. For instance,

$$(1\ 2\ 4)(3\ 5) \bullet (x_1^2 x_2 + x_5^3) = x_2^2 x_4 + x_3^3.$$

EXAMPLE 10.28. **(1)** The symmetries of a cube act on the three coordinate axises

$$x-\text{axis}, \qquad y-\text{axis}, \qquad z-\text{axis}.$$

For instance, the 120° rotation through the line connecting the origin and $(1,1,1)$ is

$$\rho\colon (x,y,z) \mapsto (z,x,y),$$

and it acts by taking the x–axis to the y–axis since $\rho((t,0,0)) = (0,t,0)$. Meanwhile the 180° rotation along the upper face, $\varrho_{\pi,\mathbf{k}}$, (i.e. through the line connecting the origin with $(0,0,1)$) is

$$\tau\colon (x,y,z) \mapsto (-x,-y,z).$$

This can be easily visualized and in doing so seen to take the x–axis to the x–axis, the y–axis to the y–axis, and the z–axis to the z–axis, i.e. it is in the kernel of this action.

In terms of the basic 90° rotations we have that (using the dictionary where the x-axis is 1, the y-axis is 2, and the z-axis is 3)

$$\varrho_{\mathbf{i},\frac{\pi}{2}} \mapsto (2\,3) \qquad \varrho_{\mathbf{j},\frac{\pi}{2}} \mapsto (1\,3) \qquad \varrho_{\mathbf{k},\frac{\pi}{2}} \mapsto (1\,2).$$

This is easily explained by noting that each of these basic rotations "flip what it is not." For instance, $\varrho_{\mathbf{i},\frac{\pi}{2}}$ flips the y and z axis. In any case we have

$$\begin{aligned} \varphi_1 &: S\left(\blacklozenge^3\right) \to \mathcal{S}_3 \\ \varrho_{\mathbf{i},\frac{\pi}{2}} &\mapsto (2\,3) \\ \varrho_{\mathbf{j},\frac{\pi}{2}} &\mapsto (1\,3) \\ \varrho_{\mathbf{k},\frac{\pi}{2}} &\mapsto (1\,2). \end{aligned}$$

which is clearly surjective. The shortest reason why this is is because $(1\,2)$ and $(2\,3)$ together generate $\mathcal{S}_3$ and these are in the image of φ_1 via $\varrho_{\mathbf{k},\frac{\pi}{2}}$ and $\varrho_{\mathbf{i},\frac{\pi}{2}}$ respectively.

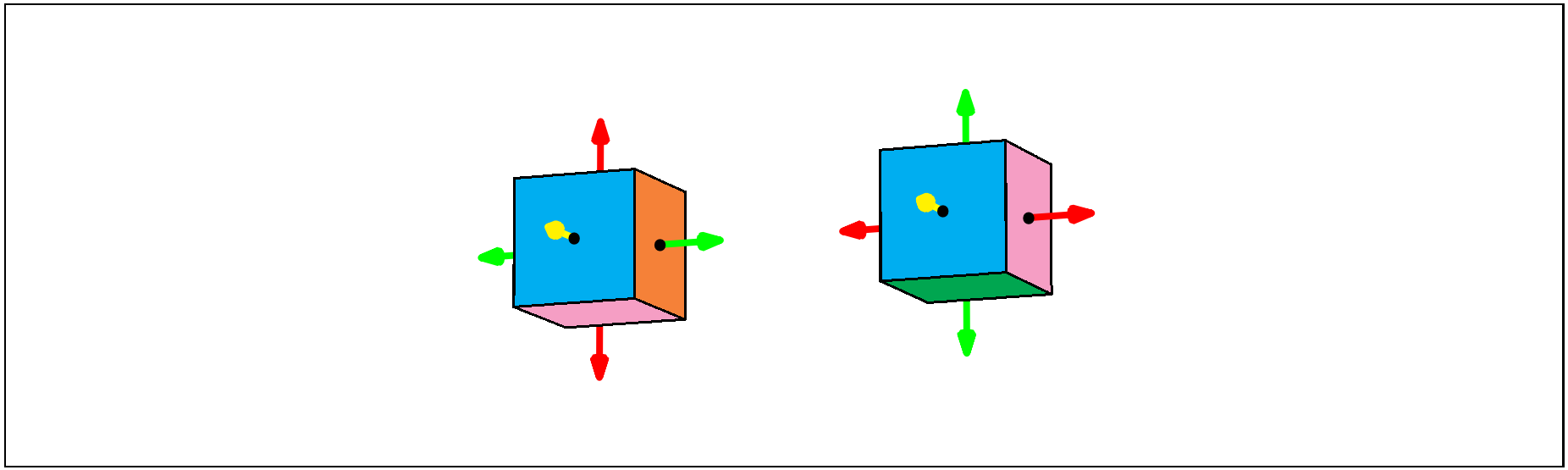
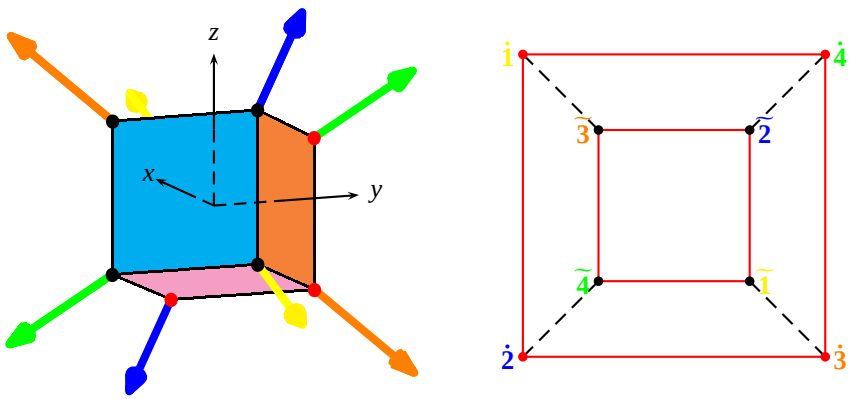


Figure 10.1: 90° rotation along the x-axis (better stated) $\varrho_{\mathbf{i},\frac{\pi}{2}}$. The x-axis is in yellow, the y-axis is in green and the z-axis is in red.

(2) The symmetries of the cube also act on the set of four diagonals which connect opposite vertices of the cube. Let's call these four diagonals $\mathbf{L}$. I claim this action, unlike the previous, has no kernel (and as such is, by definition, faithful). To see this, I flatten out the cube so we can see it better on our flat surface: (Here I drew in the diagonal connecting $\widetilde{3}$ and $\dot{3}$. While it may look like this diagonal goes through the vertex $\widetilde{1}$, this is only an illusion.) Like the symmetries of all polytopes, if a symmetry fixes all vertices of the cube it must be the identity. Therefore, if some non–identity symmetry, $\sigma \in S(\blacklozenge^3)$, acted trivially on the set $\mathbf{L}$, then since it is not the identity it must swap some vertex, v, with its opposite vertex, $\widehat{v}$. Now the three neighbors of v are strictly closer to v, then they are to $\widehat{v}$. Thus, each of the three neighbors of v must also map to their opposites. This is because symmetries of a cube preserve distances between points. Therefore, if we must have that $-\sigma$ must fix (at least) four vertices, namely v and its 3 neighbors. But either using a bit of linear algebra, or some basic geometry, we see that this implies that $-\sigma = \mathbf{I}_2$, the identity. Hence $\sigma = -\mathbf{I}_2$. But, since the determinant of $-\mathbf{I}_2$ is -1 (and not positive) it is not in $S(\blacklozenge^3)$ (it is however in $S^*(\blacklozenge^3)$). Therefore there is no non-identity element of



$S(\blacklozenge^3)$ in the kernel of the action of this group on $\mathbf{L}$. As such, we have a faithful action

$$S\left(\blacklozenge^3\right) \cup \mathbf{L}.$$

So, its permutation representation

$$S\left(\blacklozenge^3\right) \to S_{\mathbf{L}} \cong \mathcal{S}_4$$

is injective. But both groups, $S(\blacklozenge^3)$ and $\mathcal{S}_4$, have 24 elements. Therefore, the permutation representation of this action (which is always a group homomorphism), as an injective map between finite sets with the same number of elements, is a bijection. Therefore, the permutation representation is a bijective homomorphism, i.e. an isomorphism. Therefore we have proven the following rather surprising result:

THEOREM 10.29. $S\left(\blacklozenge^3\right) \cong \mathcal{S}_4$.

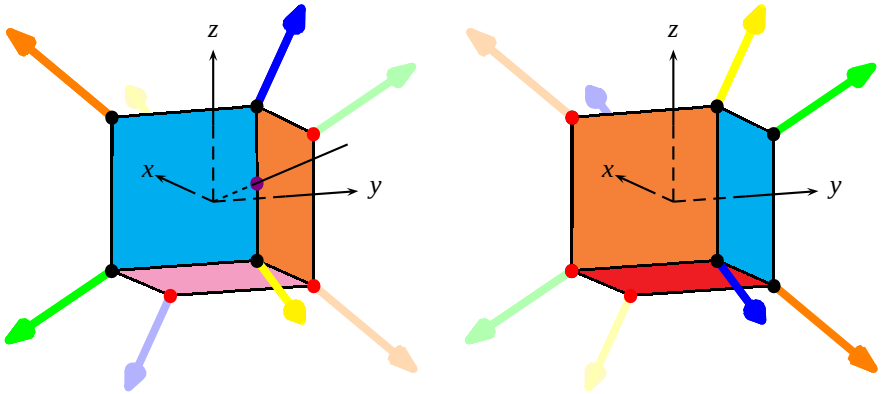
Let's provide a bit more detail: we have an isomorphism

$$\begin{aligned}
\varphi_2 : S\left(\blacklozenge^3\right) &\to \mathcal{S}_4 \\
: \varrho_{\mathbf{i},\frac{\pi}{2}} &\mapsto (\bullet\,\bullet\,\bullet\,\bullet) = (1\,2\,3\,4) \\
\varrho_{\mathbf{j},\frac{\pi}{2}} &\mapsto (\bullet\,\bullet\,\bullet\,\bullet) = (1\,3\,4\,2) \\
\varrho_{\mathbf{k},\frac{\pi}{2}} &\mapsto (\bullet\,\bullet\,\bullet\,\bullet) = (1\,3\,2\,4).
\end{aligned}$$

(3) Composing the previous two examples gives us an epimorphism:

$$\begin{aligned}
\varphi = \varphi_1 \circ \varphi_2^{-1} : \mathcal{S}_4 &\to \mathcal{S}_3 \\
: (1\,2\,3\,4)(\hookleftarrow \varrho_{\mathbf{i},\frac{\pi}{2}} \mapsto) &\mapsto (2\,3) \\
(1\,3\,4\,2)(\hookleftarrow \varrho_{\mathbf{j},\frac{\pi}{2}} \mapsto) &\mapsto (1\,3) \\
(1\,3\,2\,4)(\hookleftarrow \varrho_{\mathbf{k},\frac{\pi}{2}} \mapsto) &\mapsto (1\,2).
\end{aligned}$$

Since 4–cycles are clearly more difficult to work with than our favorite 2-cycles, the adjacent transpositions, $(1\,2)(3)(4)$, etc, let's see where these guys go. How do we realize $(1\,2)(3)(4)$ as a symmetry of the cube? Well if we rotate the cube in such a way that it fixes the mid point of the line segment connecting the relevant vertices, this would work



In terms of a rotation this is

$$\varrho_{\pi,\mathbf{i}+\mathbf{j}}$$

or in terms of a matrix this is

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

This means we have

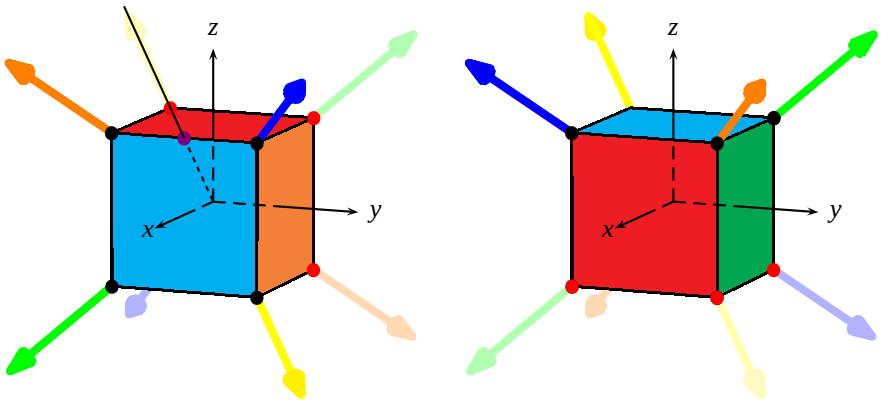
$$\varphi_2(\varrho_{\pi,\mathbf{i}+\mathbf{j}}) = (1\,2)(3)(4).$$

Looking that the matrix, or at the cube itself, we see that this rotation swaps the x and y-axes while fixing the z-axis. Putting things together we have

$$(1\,2)(3)(\hookleftarrow \varrho_{\pi,\mathbf{i}+\mathbf{j}} \hookrightarrow) \mapsto (1\,2)(3)(4).$$[VII]

What rotation realizes $(2\,3)$ under φ_2? Well

[VII] This is not a coincidence of course! I choose my dictionaries in such a way that this works out!



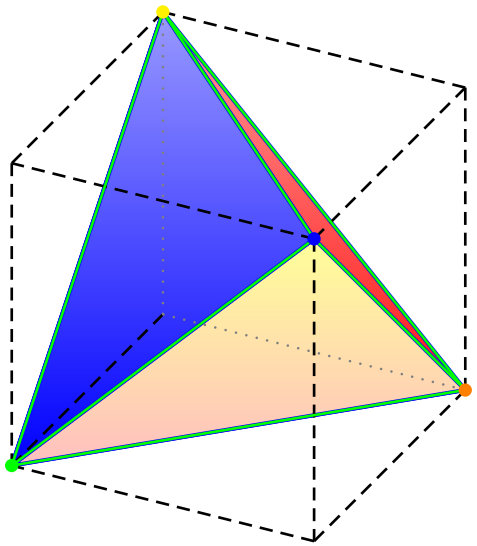
Since this swaps the x and z-axes we have

$$(1\,3)(2)(\leftarrow \varrho_{\pi,\mathbf{i+k}} \mapsto) \mapsto (2\,3)(1)(4).$$

Lastly, we could also easy to see that

$$(1\,2)(3)(\leftarrow \varrho_{\pi,\mathbf{i-j}} \mapsto) \mapsto (3\,4)(1)(2).$$

(4) Even easier than the cube, were we needed to double up the vertices, is the tetrahedron. For instance, this time there is nothing to prove in regards to the faithfulness of the action. Of course the only symmetry of a tetrahedron which fixes all vertices must be the identity, so the action is faithful.



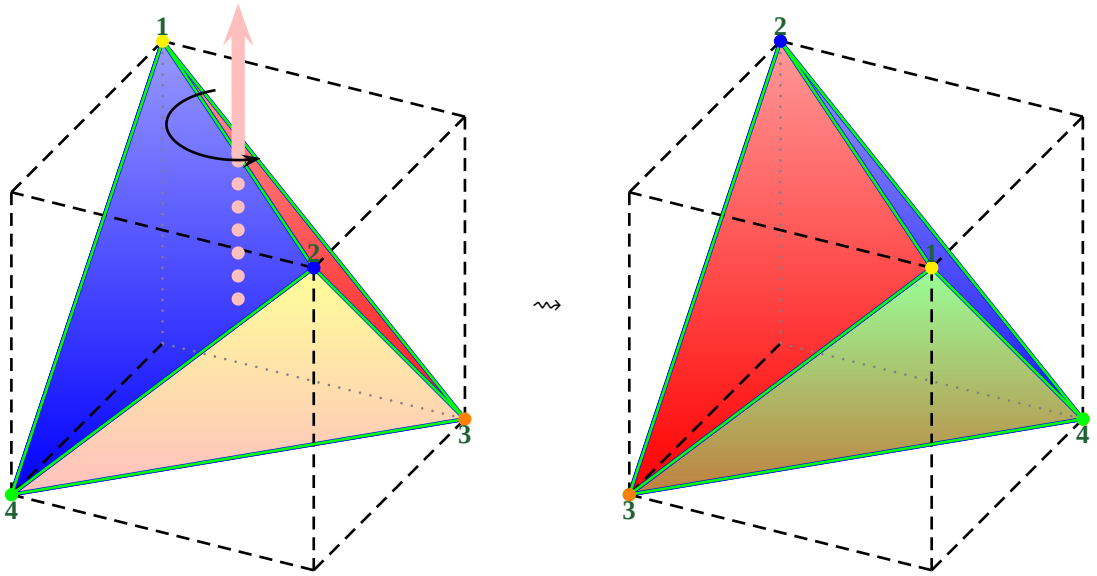
Rotation about the face with blue, orange, and green vertices (i.e. opposite the yellow vertex) induces the permutation $(\bullet)(\bullet\ \bullet\ \bullet)$. The rotation about the face opposite to the green vertex induces the permutation $(\bullet)(\bullet\ \bullet\ \bullet)$. Using our previous dictionary, we have

$$\varrho_{\frac{\pi}{3},\mathbf{i}+\mathbf{j}-\mathbf{k}} \mapsto (1)(2\,4\,3) \quad \text{and} \quad \varrho_{\frac{\pi}{3},-\mathbf{i}+\mathbf{j}+\mathbf{k}} \mapsto (1\,2\,3)(4).$$

Note that both of these permutations are even, and so $S\left(\blacktriangleright^3\right)$ is isomorphic to a subgroup of $\mathcal{A}_4$. Moreover we know that $S\left(\blacktriangleright^3\right)$ and $\mathcal{A}_4$ have the same order, therefore we have proven that

THEOREM 10.30. $S\left(\blacktriangleright^3\right) \cong \mathcal{A}_4$.

I should point out what the $\underline{2}\,\underline{2}$–cycles look like under this isomorphism. For the action on the cube they "are" 180° degree rotations about the faces. For the tetrahedron they are rotations about the edges. For example the rotation $\varrho_{\pi,\mathbf{k}}$ is



and induces the permutation $(1\,2)(3\,4)$.

EXERCISE 10.23. Describe the stabilizer (in $\mathcal{S}_4$) of the polynomial (in four variables)

$$x_1x_2 + x_3x_4$$

where the action is as defined in 4 back on page 504.

SOLUTION

10.2.1 Matrix Groups, Bilinear forms, and the Symplectic Groups

Throughout this short section, $\mathbb{F}$ is a field. Meaning that $\mathbb{F}$ is a commutative ring with every non-zero element having a multiplicative inverse. See Definition 3.142 back on page 152.

Definition 10.31.

(1) A *bilinear form* is a bilinear function, $\mathrm{B} : R^n \times R^n \to R$: i.e. a function which satisfies[VIII]

$$\begin{aligned} \mathrm{B}(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{w}) &= \mathrm{B}(\mathbf{v}_1, \mathbf{w}) + \mathrm{B}(\mathbf{v}_2, \mathbf{w}) \\ \mathrm{B}(\mathbf{v}, \mathbf{w}_1 + \mathbf{w}_2) &= \mathrm{B}(\mathbf{v}, \mathbf{w}_1) + \mathrm{B}(\mathbf{v}, \mathbf{w}_2) \\ \mathrm{B}(c\mathbf{v}, \mathbf{w}) &= c\mathrm{B}(\mathbf{v}, \mathbf{w}) = \mathrm{B}(\mathbf{v}, c\mathbf{w}). \end{aligned}$$

(2) A bilinear form is *non-degenerate* if for all $\mathbf{v} \in R^n$ there exists a $\mathbf{w} \in R^n$ with $\mathrm{B}(\mathbf{v}, \mathbf{w}) \neq 0$. It is *degenerate* otherwise.

(3) A bilinear form is *symmetric* we have

$$\mathrm{B}(\mathbf{v}, \mathbf{w}) = \mathrm{B}(\mathbf{w}, \mathbf{v})$$

and is *anti-symmetric* if

$$\mathrm{B}(\mathbf{v}, \mathbf{w}) = -\mathrm{B}(\mathbf{w}, \mathbf{v}).$$

Exercise 10.24. Define

$$\begin{aligned} \bullet : \mathbf{GL}_n(\mathbb{F}) \circlearrowright \mathbf{M}_n(\mathbb{F}) \\ : \mathbf{A} \bullet \mathbf{X} = \mathbf{A}\mathbf{X}\mathbf{A}^\top. \end{aligned}$$

SOLUTION

(1) Prove that $\bullet$ is an action.

(2) When $\mathbb{F} = \mathbb{R}$, the real numbers, what is the name of the familiar group that the stabilizer of the identity matrix, $\mathbf{I}_n$, equal to? In symbols:

$$\mathbf{GL}_n(R)_{\mathbf{I}_n} = ?.$$

(3) A matrix $\mathbf{X} \in \mathbf{M}_n(R)$ is *symmetric* if $\mathbf{X}^\top = \mathbf{X}$. Prove that $\bullet$ gives us an action on the set of symmetric matrices by showing that the action preserves the set of symmetric matrices. (That is show that if $\mathbf{X}$ is symmetric, then so is $\mathbf{A} \bullet \mathbf{X}$.)

(4) A symmetric matrix, $\mathbf{X} \in \mathbf{M}_n(R)$, is called *positive definite* (or pd) if for non–zero vectors $\mathbf{v} \in \mathbb{R}^n$ (thought of as a column vector) we have that

$$\mathbf{v}^\top \cdot \mathbf{X} \cdot \mathbf{v} > 0.$$[IX]

Prove that $\bullet$ gives us an action on the set of pd matrices by showing that the action preserves the set of p.d. matrices. (That is show that if $\mathbf{X}$ is pd, then so is $\mathbf{A} \bullet \mathbf{X}$.)

[VIII] All vectors and scalars in this definition are arbitrary.

For a square matrix $\mathbf{X} \in \mathbf{M}_n(\mathbb{F})$ define $\widetilde{\mathbf{X}}\colon \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$[X] via

$$\widetilde{\mathbf{X}}(\mathbf{v}, \mathbf{w}) := \mathbf{v}^\top \cdot \mathbf{X} \cdot \mathbf{w}.$$

Note that $\widetilde{\mathbf{X}}$ is a bilinear form. For instance:

$$\begin{aligned}
\widetilde{\mathbf{X}}(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{w}) &= (\mathbf{v}_1 + \mathbf{v}_2)^\top \cdot \mathbf{X} \cdot \mathbf{w} \\
&= (\mathbf{v}_1^\top + \mathbf{v}_2^\top) \cdot \mathbf{X} \cdot \mathbf{w} \\
&= (\mathbf{v}_1^\top \cdot \mathbf{X} + \mathbf{v}_2^\top \cdot \mathbf{X}) \cdot \mathbf{w} \\
&= \mathbf{v}_1^\top \cdot \mathbf{X} \cdot \mathbf{w} + \mathbf{v}_2^\top \cdot \mathbf{X} \cdot \mathbf{w} \\
&= \widetilde{\mathbf{X}}(\mathbf{v}_1, \mathbf{w}) + \widetilde{\mathbf{X}}(\mathbf{v}_2, \mathbf{w}).
\end{aligned}$$

Checking that

$$\widetilde{\mathbf{X}}(\mathbf{v}, \mathbf{w}_1 + \mathbf{w}_2) = \widetilde{\mathbf{X}}(\mathbf{v}, \mathbf{w}_1) + \widetilde{\mathbf{X}}(\mathbf{v}, \mathbf{w}_2) \quad \text{and} \quad \widetilde{\mathbf{X}}(c\mathbf{v}, \mathbf{w}) = c\widetilde{\mathbf{X}}(\mathbf{v}, \mathbf{w}) = \widetilde{\mathbf{X}}(\mathbf{v}, c\mathbf{w})$$

are just as easy.

Conversely, for any bilinear form $\mathrm{B} : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ there is an associated n by n square matrix, $\widehat{\mathrm{B}}$, whose i, j entry is

$$\widehat{\mathrm{B}}(i, j) \stackrel{\text{def}}{=} \mathrm{B}(\mathbf{e}_i, \mathbf{e}_j).$$

Exercise 10.25. Prove that $\mathbf{A} \mapsto \widetilde{\mathbf{A}}$ and $\mathrm{B} \mapsto \widehat{\mathrm{B}}$ are inverse processes.

Solution

Exercise 10.26. Prove

(1) Prove that $\widetilde{\mathbf{X}}$ is non-degenerate iff $\mathbf{X}$ is an invertible matrix.

(2) Prove that $\widetilde{\mathbf{X}}$ is symmetric iff $\mathbf{X}$ is a symmetric matrix.

(3) Prove that $\widetilde{\mathbf{X}}$ is anti–symmetric iff $\mathbf{X}$ is an anti–symmetric matrix.

(4) Prove TFAE

- $\mathbf{A}$ fixes $\mathbf{X}$ under the action from the previous exercise, $\mathbf{A}\mathbf{X}\mathbf{A}^\top$,
- For all $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{F}^n$,
$$\widetilde{\mathbf{X}}(\mathbf{v}, \mathbf{w}) = \widetilde{\mathbf{X}}(\mathbf{A}(\mathbf{v}), \mathbf{A}(\mathbf{w})).$$
- For all i and j
$$\widetilde{\mathbf{X}}(\mathbf{e}_i, \mathbf{e}_j) = \widetilde{\mathbf{X}}(\mathbf{A}(\mathbf{e}_i), \mathbf{A}(\mathbf{e}_j)).$$

Hint! Compare this with the proof of Theorem C.6 coming up on page 759 in the appendix.

Solution

[IX] This only makes sense for fields contained in $\mathbb{R}$ since we are talking about $>$, which makes no sense over $\mathbb{C}$ or $\mathbb{Z}_n$!

[X] As usual, $\mathbb{F}^n$ is thought of as column vectors.

Define the 2 by 2 matrix

$$\mathbf{J}_2 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

More generally, define $\mathbf{J}_{2n}$ as the 2n by 2n block matrix with $\mathbf{J}_2$'s down the diagonal

$$\mathbf{J}_{2n} := \mathrm{D}(\underbrace{\mathbf{J}_2, \mathbf{J}_2, \ldots, \mathbf{J}_2}_{n-\text{times}}).$$

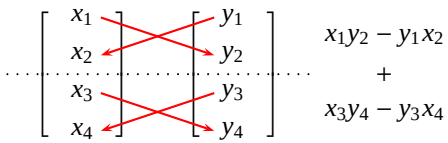
For instance, for $n = 2$, we have that

$$\mathbf{J}_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

and for $n = 3$ we have

$$\mathbf{J}_6 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}.$$

It is easy to picture how to calculate $\widetilde{\mathbf{J}}_{2n}$ via taking n 2 by 2 determinants and adding them together. For instance, (in the case that $n = 4$)



$$x_1y_2 - y_1x_2 + x_3y_4 - y_3x_4$$

$$(x_1y_2 - y_1x_2) + (x_3y_4 - y_3x_4)$$

The stabilizer of $\mathbf{J}_{2n}$ under the action $\bullet$ is denoted $\mathbf{Sp}_{2n}(\mathbb{F})$.[XI] Said another way, it's the set of matrices $\mathbf{A}$ which satisfy

$$\widetilde{\mathbf{J}}_{2n}(\mathbf{v}, \mathbf{w}) = \widetilde{\mathbf{J}}_{2n}(\mathbf{A}(\mathbf{v}), \mathbf{A}(\mathbf{w}))$$

for all $2n$–vectors $\mathbf{v}$ and $\mathbf{w}$. It is called the *symplectic group* (of degree n and $\mathbb{F}$ coefficients) and matrices in this group are called *symplectic*.

[XI] So

$$\mathbf{Sp}_{2n}(\mathbb{F}) \stackrel{\text{def}}{=} \{\mathbf{A} \in \mathbf{M}_{2n}(\mathbb{F}) : \mathbf{A}\mathbf{J}_{2n}\mathbf{A}^\top = \mathbf{J}_{2n}\}.$$

EXAMPLE 10.32.

(1) The matrix

$$\mathbf{A} := \begin{pmatrix} 3 & 1 & -2 & 1 \\ 0 & 5 & 6 & -1 \\ 1 & 0 & 4 & 1 \\ 0 & 1 & 6 & 1 \end{pmatrix}$$

is not in $\mathbf{Sp}_4(\mathbb{R})$ since

$$\widetilde{\mathbf{J}}_4(\mathbf{e}_1, \mathbf{e}_2) = 1$$

but

$$\begin{aligned} \widetilde{\mathbf{J}}_4(\mathbf{A}(\mathbf{e}_1), \mathbf{A}(\mathbf{e}_2)) &= \widetilde{\mathbf{J}}_4(3\mathbf{e}_1 + \mathbf{e}_3, \mathbf{e}_1 + 5\mathbf{e}_2 + \mathbf{e}_4) \\ &= 3\widetilde{\mathbf{J}}_4(\mathbf{e}_1, \mathbf{e}_1) + 15\widetilde{\mathbf{J}}_4(\mathbf{e}_1, \mathbf{e}_2) + 3\widetilde{\mathbf{J}}_4(\mathbf{e}_1, \mathbf{e}_4) \\ &\quad + \widetilde{\mathbf{J}}_4(\mathbf{e}_3, \mathbf{e}_1) + 5\widetilde{\mathbf{J}}_4(\mathbf{e}_3, \mathbf{e}_2) + \widetilde{\mathbf{J}}_4(\mathbf{e}_3, \mathbf{e}_4) \\ &= 15 + 1 \\ &= 16. \end{aligned}$$

Since $16 \neq 1$ the matrix $\mathbf{A}$ is not symplectic.

(2) On the other hand, the matrix

$$\mathbf{A} := \begin{pmatrix} 0 & 1 & 1 & 1 \\ -1 & 2 & 1 & 1 \\ -1 & 1 & 2 & 1 \\ 1 & -1 & -1 & 0 \end{pmatrix}$$

is symplectic. This can be checked directly by computing

$$\mathbf{A} \cdot \mathbf{J}_4 \mathbf{A}^\top = \begin{pmatrix} 0 & 1 & 1 & 1 \\ -1 & 2 & 1 & 1 \\ -1 & 1 & 2 & 1 \\ 1 & -1 & -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 1 \\ -1 & 2 & 1 & 1 \\ -1 & 1 & 2 & 1 \\ 1 & -1 & -1 & 0 \end{pmatrix}^\top$$

and showing it to equal $\mathbf{J}_4$. For more practice computing $\mathbf{J}$'s, we verify this in some special cases. For instance

$$\begin{aligned} \widetilde{\mathbf{J}}_4(\mathbf{A}(\mathbf{e}_1), \mathbf{A}(\mathbf{e}_2)) &= \widetilde{\mathbf{J}}_4(-\mathbf{e}_2 - \mathbf{e}_3 + \mathbf{e}_4, \mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3 - \mathbf{e}_4) \\ &= \widetilde{\mathbf{J}}_4(-\mathbf{e}_2, \mathbf{e}_1 + 2\mathbf{e}_2) + \widetilde{\mathbf{J}}_4(-\mathbf{e}_3 + \mathbf{e}_4, \mathbf{e}_3 - \mathbf{e}_4) \\ &= \widetilde{\mathbf{J}}_4(-\mathbf{e}_2, \mathbf{e}_1 + 2\mathbf{e}_2) - \widetilde{\mathbf{J}}_4(\mathbf{e}_3 - \mathbf{e}_4, \mathbf{e}_3 - \mathbf{e}_4) \qquad \text{(For any } \mathbf{v}, \widetilde{\mathbf{J}}(\mathbf{v}, \mathbf{v}) = 0) \\ &= \widetilde{\mathbf{J}}_4(-\mathbf{e}_2, \mathbf{e}_1) - 2\widetilde{\mathbf{J}}_4(\mathbf{e}_2, 2\mathbf{e}_2) \qquad \text{(For any } \mathbf{v}, \widetilde{\mathbf{J}}(\mathbf{v}, \mathbf{v}) = 0) \\ &= \widetilde{\mathbf{J}}_4(\mathbf{e}_1, \mathbf{e}_2). \end{aligned}$$

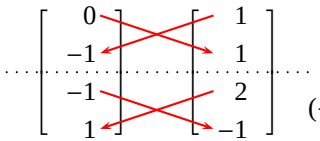
Since we get

$$\widetilde{\mathbf{J}}_4(\mathbf{A}(\mathbf{e}_1), \mathbf{A}(\mathbf{e}_1)) = \widetilde{\mathbf{J}}_4(\mathbf{A}(\mathbf{e}_2), \mathbf{A}(\mathbf{e}_2)) = \widetilde{\mathbf{J}}_4(\mathbf{A}(\mathbf{e}_3), \mathbf{A}(\mathbf{e}_3)) = \widetilde{\mathbf{J}}_4(\mathbf{A}(\mathbf{e}_4), \mathbf{A}(\mathbf{e}_4)) = 0$$

for free we still to perform 5 more calculations:

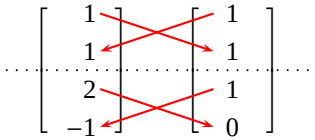
$$\mathbf{J}(\mathbf{e}_1, \mathbf{e}_3) = 0 \qquad \mathbf{J}(\mathbf{e}_1, \mathbf{e}_4) = 0$$
$$\mathbf{J}(\mathbf{e}_2, \mathbf{e}_3) = 0 \qquad \mathbf{J}(\mathbf{e}_2, \mathbf{e}_4) = 0$$
$$\mathbf{J}(\mathbf{e}_3, \mathbf{e}_4) = 1.$$

This may sound overwhelming, but in reality you should be able to quickly eyeball these calculations. I show the first and last calculations only



$$(0)\cdot(1) - (1)\cdot(-1) = 1$$
$$+$$
$$(-1)\cdot(-1) - (2)\cdot(-1) = -1$$

$$1 + (-1) = 0$$



$$(1)\cdot(1) - (1)\cdot(1) = 0$$
$$+$$
$$(2)\cdot(0) - (1)\cdot(-1) = 1$$

$$0 + 1 = 1$$

EXERCISE 10.27. Let $\mathbb{F}$ be a field. [XII]

(1) State a theorem given earlier in this chapter which proves that $\mathbf{Sp}_{2n}(\mathbb{F})$ is a group.

SOLUTION

(2) Prove that $\mathbf{Sp}_2(\mathbb{F}) = \mathbf{SL}_2(\mathbb{F})$.

(3) Prove that, in general, $\mathbf{Sp}_4(\mathbb{F}) \neq \mathbf{SL}_4(\mathbb{F})$.

(4) Wiki something and report.

NOTATION 10.33. Fix $\mathbf{w} \in \mathbb{F}^{2n}$ and $\lambda \in \mathbb{F}$. Define

$$t_{\mathbf{w}}^{\lambda}(\mathbf{v}) := \mathbf{v} + \lambda \widetilde{\mathbf{J}}(\mathbf{v}, \mathbf{w})\mathbf{w}.$$

Also, we denote $t_{\mathbf{v}}^{1}$ as $t_{\mathbf{v}}$.

EXERCISE 10.28.

(1) Prove that $t_{\mathbf{w}}^{\lambda} \in \mathbf{Sp}_{2n}(\mathbb{F})$.
Hint! Expand

$$\widetilde{\mathbf{J}}(t_{\mathbf{w}}^{\lambda}(\mathbf{v}), t_{\mathbf{w}}^{\lambda}(\mathbf{v}'))$$

by using bilinearity, and skew-symmetry of $\mathbf{J}$, and show that it reduces back down to $\widetilde{\mathbf{J}}(\mathbf{v}, \mathbf{v}')$.

(2) Prove for $\lambda, \mu \in R$ that

$$t_{\mathbf{w}}^{\lambda} \circ t_{\mathbf{w}}^{\mu} = t_{\mathbf{w}}^{\lambda+\mu} = t_{\mathbf{w}}^{\mu} \circ t_{\mathbf{w}}^{\lambda}.$$

(3) Prove that $t_{\mathbf{w}}^{\lambda}$ and $t_{\mathbf{v}}^{\mu}$ commute iff $\mathbf{v}$ and $\mathbf{w}$ are $\mathbf{J}$–orthogonal (meaning $\widetilde{\mathbf{J}}(\mathbf{v}, \mathbf{w}) = 0$.)

(4) Prove that for any $\mathbf{v} \in \mathbb{F}^{2n}$ there exists a $\widehat{\mathbf{v}}$ so that

$$\widetilde{\mathbf{J}}(\mathbf{v}, \widehat{\mathbf{v}}) = 1.$$

(5) Prove that if $\widetilde{\mathbf{J}}(\mathbf{v}, \mathbf{w}) = \lambda \neq 0$, then

$$t_{\mathbf{w}-\mathbf{v}}^{\lambda^{-1}}(\mathbf{v}) = \mathbf{w}.$$

(6) Prove that for any $\mathbf{v}$, $\mathbf{w}$ in $\mathbb{F}^{2n}$ which are $\mathbf{J}$–orthogonal, prove that there exists $\mathbf{u} \in \mathbb{F}^{2n}$ with

$$\widetilde{\mathbf{J}}(\mathbf{v}, \mathbf{u}) \neq 0 \quad \text{and} \quad \widetilde{\mathbf{J}}(\mathbf{w}, \mathbf{u}) \neq 0.$$

Hint! Play with several cases of the $\widehat{\mathbf{v}}'s$ from a few items back.

- If $\widehat{\mathbf{v}}$ is not $\mathbf{J}$–orthogonal to $\mathbf{w}$, define $\mathbf{u}$ to be $\widehat{\mathbf{v}}$. If it is orthogonal, then move onto the next step.

SOLUTION

[XII] Most parts of these next two exercises work even if $\mathbb{F}$ is just a plain ring. A couple need to assume that R is commutative, and the last few parts need to require that $\mathbb{F}$ is a field. Which ones?

- If $\widehat{\mathbf{w}}$ is not $\mathbf{J}$–orthogonal to $\mathbf{v}$, define $\mathbf{u}$ to be $\widehat{\mathbf{w}}$. If it is orthogonal, then move onto the next step.
- If both $\widetilde{\mathbf{J}}(\widehat{\mathbf{v}}, \mathbf{w}) = \widetilde{\mathbf{J}}(\widehat{\mathbf{w}}, \mathbf{v})$, then define $\mathbf{u} = \widehat{\mathbf{v}} + \widehat{\mathbf{w}}$.

(7) Prove that the action of $\mathbf{Sp}_{2n}(\mathbb{F}) \circlearrowright \mathbb{F}^{2n}$ is transitive.
Hint! There are two cases. For $\mathbf{v}$ and $\mathbf{w} \in \mathbb{F}^{2n}$ if $\mathbf{v}$ and $\mathbf{w}$ are not $\mathbf{J}$–orthogonal, then use Part **(5)** to find an element of $\mathbf{Sp}_{2n}(\mathbb{F})$ which takes $\mathbf{v}$ to $\mathbf{w}$. If they are $\mathbf{J}$–orthogonal, then first use Part **(6)** before using Part **(5)** twice.

(8) Prove that for $n \geq 1$, that $\mathbf{Sp}_{2n}(\mathbb{F}) \circlearrowright \mathbb{F}^{2n}$ is not 2–transitive.
Hint! You can't take a $\mathbf{J}$–orthogonal pair to a non $\mathbf{J}$–orthogonal pair.

EXERCISE 10.29.

SOLUTION

(1) For a symplectic matrix $\mathbf{A}$ prove that

$$\mathbf{A} \cdot t_{\mathbf{v}}^{\lambda} \cdot \mathbf{A}^{-1} = t_{\mathbf{A}\cdot\mathbf{v}}^{\lambda}.$$

Careful! This is false if $\mathbf{A}$ is not symplectic, so be sure to use the fact that $\mathbf{A}$ is symplectic!

(2) Prove

$$t_{\mathbf{v}}^{\lambda} = t_{\mathbf{w}}^{\lambda} \Leftrightarrow \mathbf{v} = \pm\mathbf{w}.$$

(3) Prove that a symplectic matrix $\mathbf{A}$ commutes with $t_{\mathbf{v}}^{\lambda}$ iff $\mathbf{A} \cdot \mathbf{v} = \pm\mathbf{v}$.

(4) Using the previous parts of this exercise prove that the center of $\mathbf{Sp}_{2n}(\mathbb{F})$ consists linear transformations that map every vector to itself or to negative itself.
Hint! Use previous part while considering the $t_{\mathbf{v}}$'s.

(5) Prove that $\mathcal{C}\mathbf{Sp}_{2n}(\mathbb{F})$ contains just $\pm\mathbf{I}_{2n}$.
Hint! We know that for all $\mathbf{v}$ we have $\mathbf{A}(\mathbf{v}) = \pm\mathbf{v}$. Be careful, you still need to show that we can't have for some $\mathbf{v}$ that $\mathbf{A}(\mathbf{v}) = -\mathbf{v}$ and for some other $\mathbf{w}$ that $\mathbf{A}(\mathbf{w}) = \mathbf{w}$.

EXERCISE 10.30. Prove that the action

SOLUTION

(1) $\mathbf{SO}_3(\mathbb{R}) \circlearrowright \mathbb{R}^3$ is not transitive.

(2) $\mathbf{SO}_3(\mathbb{R}) \circlearrowright \mathbb{S}^2$ is transitive.

(3) $\mathbf{SO}_3(\mathbb{R}) \circlearrowright \mathbb{S}^2$ is not 2-transitive.

Define the *Stiefel manifold*, $\mathbf{V}_2(\mathbb{R}^3)$, as the set of orthonormal pairs. That is

$$\mathbf{V}_2(\mathbb{R}^3) \stackrel{\text{def}}{=} \{(\widehat{\mathbf{u}}, \widehat{\mathbf{v}}) \in \mathbb{S}^2 \times \mathbb{S}^2 : \widehat{\mathbf{u}} \bullet \widehat{\mathbf{v}} = 0\}.$$ [XIII]

(4) Prove that

$$\mathbf{A} \star (\widehat{\mathbf{u}}, \widehat{\mathbf{u}'}) = (\mathbf{A}(\widehat{\mathbf{u}}), \mathbf{A}(\widehat{\mathbf{u}'}))$$

defines an action $\mathbf{O}_3(\mathbb{R}) \circlearrowright \mathbf{V}_2(\mathbb{R}^3)$.

(5) Find the stabilizer (in $\mathbf{O}_3(\mathbb{R})$) of $(\mathbf{e}_1, \mathbf{e}_2) \in \mathbf{V}_2(\mathbb{R}^3)$ under this action.

(6) Prove that $\mathbf{SO}_3(\mathbb{R})$ acts freely and transitively on $\mathbf{V}_2(\mathbb{R}^3)$.

Here $\bullet$ is the dot product of vectors, not an action!

10.3 A Bit More Theory

Since conjugate subgroups are isomorphic, a direct consequence of the first part of the next theorem is that for two points, x and y, in the same orbit of an action of G

$$G_x \cong G_y.$$

In particular, the same number of elements of G stabilize any two points in the same orbit of an action. As such, in a transitive action any two points in X are stabilizer by the same number of elements of G.

Theorem 10.34. *For subsets $A \subseteq X$ and $S \subseteq G$ we have*

(1) $G_{g \bullet A} = g\, G_A\, g^{-1}$

(2) $g \bullet X^S = X^{gSg^{-1}}$.

Proof.

(1) We show the equality of the two sets by following a straight forward follow your nose approach.

$$\begin{aligned}
\overline{g} \in G_{g \bullet A} &\Leftrightarrow \forall a \in A\ \overline{g} \bullet (g \bullet a) \in g \bullet A \\
&\Leftrightarrow \forall a \in A,\ \exists \overline{a} \in A \text{ s.t. } \overline{g} \bullet (g \bullet a) = g \bullet \overline{a} \\
&\Leftrightarrow \forall a \in A,\ \exists \overline{a} \in A \text{ s.t. } \overline{g}g \bullet a = g \bullet \overline{a} \\
&\Leftrightarrow \forall a \in A,\ \exists \overline{a} \in A \text{ s.t. } g^{-1}\overline{g}g \bullet a = \overline{a} \\
&\Leftrightarrow \forall a \in A\ (g^{-1}\overline{g}g) \bullet a \in A \\
&\Leftrightarrow g^{-1}\overline{g}g \in G_A \\
&\Leftrightarrow \overline{g} \in g\, G_A\, g^{-1}.
\end{aligned}$$

(2) The proof is similar to the proof of 1, but I'll do it out because I'm a loser.

$$\begin{aligned} x \in X^{gSg^{-1}} &\Leftrightarrow \forall s \in S,\ gsg^{-1} \bullet x = x \\ &\Leftrightarrow \forall s \in S,\ sg^{-1} \bullet x = g^{-1} \bullet x \\ &\Leftrightarrow \forall s \in S,\ s \bullet \left(g^{-1} \bullet x\right) = g^{-1} \bullet x \\ &\Leftrightarrow \forall s \in S,\ g^{-1} \bullet x \in X^s \\ &\Leftrightarrow g^{-1} \bullet x \in X^S \qquad \left(X^S = \bigcap_{s \in S} X^s\right) \\ &\Leftrightarrow x \in g \bullet X^S. \end{aligned}$$

■

Notice that Theorem 7.25 is really a special case of this theorem. As pointed out just before its statement, this theorem says, among other things the following:

COROLLARY 10.35. *If G acts transitively on X, then the stabilizer subgroups any two points in X are conjugate* (*hence isomorphic*).

PROOF. Suppose that G acts transitively on X. Then, by definition, for any x and y in X there exists a $g \in G$ so that $g \bullet x = y$. But by the theorem, this implies that

$$G_y = G_{g \bullet x} = g\, G_x\, g^{-1}.$$

Since x and y are arbitrary elements of X, the stabilizer of any two points are conjugate. ■

For instance, for the action $\mathcal{S}_4 \circlearrowright \{1, 2, 3, 4\}$, the stabilizer of the letter 4 is (the natural copy of) $\mathcal{S}_3$. However, intuitively, nothing in a permutation group should depend on what order or what you call your letters. For instance, if we had four identical balls, and we permuted three of them while holding one of them fixed, it certainly shouldn't matter which one we kept fixed. Therefore it should be intuitive in this case that the stabilizer of the letter 2, $(\mathcal{S}_4)_2$ or the stabilizer of the letter 3, $(\mathcal{S}_4)_3$, etc, could just as easily be identified with $\mathcal{S}_3$ instead of $(\mathcal{S}_4)_1$.

THEOREM 10.36. *Suppose G acts transitively on a collection of subsets of X, Λ. Then*

(1) *All elements of Λ have the same cardinality.*

(2) *If the collection Λ partitions X, then*

$$\#X = (\#\Lambda) \cdot (\#A)$$

where A is any element of Λ.

PROOF.

(1) If $A, B \in \Lambda$, then there is a $g \in G$ with $g \bullet A = B$. This is a bijection from A to B since it has an inverse $g^{-1}\bullet$.

(2) In general, if Λ is a partition of a set X we have that

$$\#X = \sum_{A \in \Lambda} \#A.$$

The first part says that every term inside of the sum has the same order, so

$$\#X = \sum_{A \in \Lambda} \#A = \#A \sum_{A \in \Lambda} 1 = \#A \#\Lambda.$$

■

The hypothesis of the theorem holds iff G preserves some equivalence relation $\sim$ defined on X in the sense that

$$x \sim y \Rightarrow (g \bullet x) \sim (g \bullet y).$$

In this case the collection Λ is the collection of equivalence classes of $X/\sim$ (don't forget that equivalence classes are subsets of the original space! Therefore $X/\sim$ is a subset of the powerset of X, $\mathcal{P}(X)$).

Example 10.37. Suppose that G acts on X, and we have $H \trianglelefteq G$. Then we claim that G preserves the orbits of H of H's action on X. In terms of the equivalence relation we are looking at the relation $\sim$ defined by

$$x \sim y \Leftrightarrow \exists\, h \in H \text{ s.t. } h \bullet x = y.$$

Then G preserves this equivalence relation since if $x \sim y$ then

$$\begin{aligned}
&\exists h \in H \text{ s.t. } h \bullet x = y \\
&\forall g \in G\ ghg^{-1} \bullet x = g \bullet y \\
&\forall g \in G\ ghg^{-1}g \bullet x = g \bullet y \\
&\Rightarrow \overline{h}g \bullet x = g \bullet y \qquad \text{(Since } H \trianglelefteq G,\ ghg^{-1} = \overline{h} \in H\text{)} \\
&\Rightarrow \exists \overline{h} \in H \text{ s.t. } \overline{h} \bullet (g \bullet x) = g \bullet y \\
&\Rightarrow g \bullet x \sim g \bullet y.
\end{aligned}$$

This means that we get an action of G on the set of H orbits, X/H with $g \in G$ taking

$$H \bullet x \qquad \text{to} \qquad H \bullet (g \bullet x).$$

Moreover, we should point out that H acts trivially in this new action. This is because, by definition, elements of H simply permute these orbits within themselves (and do not allow them to get mixed up).[I]

Exercise 10.31. Let

$$G := \left\langle (1\ \overset{x}{2}\ 3\ 4), (\overset{y}{2}\ 4) \right\rangle$$

which is isomorphic to $\mathcal{D}_4$. Show that G breaks up the orbits of $\langle y \rangle$ in the sense that there are orbits C and an element $g \in G$ s.t. $g \bullet C$ is not an orbit $\langle y \rangle$.

Solution

Example 10.38. Consider the group G from Example 9.35((6)) back on page 455. In that example we saw that the group N (as defined in that example) was normal in G. The orbit of N on the set $\overline{12}$ has 4 orbits given by

$$\dot{1} := N \bullet 1 = \{1, 9, 5\} \qquad \dot{2} := N \bullet 2 = \{2, 10, 6\}$$
$$\dot{3} := N \bullet 3 = \{3, 7, 11\} \qquad \dot{4} := N \bullet 4 = \{4, 8, 12\}.$$

[I] Think of stacked shuffle board chips getting swapped around. The stacks themselves are determined by the subgroup H and this example shows that G does not cause the stacks to be broken up.

Of course x acts on $\overline{12}/_N$ trivially (since $x \in N$). But

$$y \bullet \dot{1} = \{10, 6, 2\} = \dot{2} \qquad y \bullet \dot{2} = \dot{1} \qquad y \bullet \dot{3} = \{12, 4, 8\} = \dot{4}$$
$$y \bullet \dot{4} = \dot{3} \qquad z \bullet \dot{1} = \dot{3} \qquad z \bullet \dot{2} = \dot{4}$$
$$z \bullet \dot{3} = \dot{1} \qquad z \bullet \dot{4} = \dot{2}.$$

Therefore, we get a homomorphism $G \to \mathcal{S}_4$

$$x \mapsto (\dot{1}) \qquad y \mapsto \left(\dot{1}\ \dot{2}\right)\left(\dot{3}\ \dot{4}\right) \qquad z \mapsto \left(\dot{1}\ \dot{3}\right)\left(\dot{2}\ \dot{4}\right)$$

(which has image $V \cong \mathbb{Z}_2 \times \mathbb{Z}_2$).

10.4 Cosets

With the Cayley action of G on itself, which we defined as

$$g \bullet x = gx$$

in Theorem 10.13, are there any obvious preserved collections?

Definition 10.39. Let H be a subgroup of a group G and let $g \in G$. We define the (left) *coset* of H by $a \in G$ as the set

$$aH \stackrel{\text{def}}{=} \{ah : h \in H\} \tag{5}$$

If we want to make the operation in G more explicit, this is written $g * H$. For instance, in $\mathbb{Z}$ and $\mathbb{Z}_n$ we are looking at the sets $a + H = \{a + h : h \in H\}$. An alternate, more passive, definition is

Definition 10.40.

$$aH \stackrel{\text{def}}{=} \{\, b : b = ah \text{ for some } h \in H\}. \tag{6}$$

For instance, by closure of H we see that for any $h \in H$ that $hH = H$.[1] In fact the converse is also true. In other words, if $H = gH$, then we must have that $g \in H$. This is because if $H = hH$, then since $e_G \in H$ it is also in gH. That means there is an $h \in H$ so that $e = gh$. But this implies that $g = h^{-1}$ which is in H since H is a subgroup, and as such, closed under inversion. Thus we have proven:

Proposition 10.41. $gH = H \Leftrightarrow g \in H$. ∎

[1] Actually, it's only 100% obvious that $hH \subseteq H$ the other inclusion takes an extra second of thought ...

WARNING! 10.42. The subgroups $n\mathbb{Z} = \{nz : z \in \mathbb{Z}\}$ of $\mathbb{Z}$ are **not** (non-trivial) cosets (even though they look like them)! This is because the operation in $\mathbb{Z}$ is addition, **not** multiplication! Therefore, for a fixed integer n, the cosets of the subgroup $n\mathbb{Z}$ are our old friends

$$a + n\mathbb{Z}$$

from way back on page 75.

In fact, for a fixed subgroup $H \leqslant G$, the only coset of H which is a subgroup of G is the coset $e_G H = H$ itself! This is because (as we just showed) the coset gH can only contain the identity (which is something that is required if gH is to be a subgroup of G) in the special case that $g \in G$!

EXAMPLE 10.43.

(1) Let G denote $\mathcal{D}_4$ and let's denote the subgroup of $\mathcal{D}_4$, $\langle \rho \rangle = \{\mathbf{1}, \rho\}$, by H. We have

$$\begin{aligned} \mathbf{1}H &= \{\mathbf{1}, \mathbf{r}\} & \mathbf{r}H &= \{\mathbf{r}, \mathbf{1}\} \\ \rho H &= \{\rho, \mathbf{r}\rho^3\} & \rho\mathbf{r}H &= \{\rho\mathbf{r}, \rho\} \\ \rho^2 H &= \{\rho^2, \mathbf{r}\rho^2\} & \rho^2\mathbf{r}H &= \{\rho^2\mathbf{r}, \rho^2\} \\ \rho^3 H &= \{\rho^3, \rho^3\mathbf{r}\} & \rho^3\mathbf{r}H &= \{\rho^3\mathbf{r}, \rho^3\}. \end{aligned}$$

Notice that any two cosets in the same row are equal to each other, and there are a total of $8/2 = 4$ distinct cosets. Moreover notice that the cosets partition the bigger group, here $\mathcal{D}_4$, in the sense that every element of $\mathcal{D}_4$ is in **exactly one** coset.

(2) Denote

$$\langle (1\,2\,3) \} = \{(1), (1\,2\,3), (1\,3\,2)\rangle$$

by H and $\mathcal{A}_4$ as G. We have

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

Notice that any two cosets in the same row are equal to each other, and there are a total of $12/3 = 4$ distinct cosets. Moreover notice that, once again, the cosets partition the bigger group, here $\mathcal{A}_4$, in the sense that every element of $\mathcal{A}_4$ is in **exactly one** coset.

(3) Let's do a more difficult example (whose goal is to make us appreciate the theory to come!!):

$$H = \left\langle -\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k} \right\rangle = \left\{1, -\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}, -\frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k}\right\}$$

inside the binary tetrahedral group $B\left(\blacktriangleright^3\right)$ (see page 387). (we do not write repeats)

$$
\begin{aligned}
H &= \left\{ \quad 1, \quad -\frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k}, \quad -\frac{1}{2}-\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k} \right\} \\
-H &= \left\{ -1, \quad \frac{1}{2}-\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k}, \quad \frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k} \right\} \\
\mathbf{i}H &= \left\{ \quad \mathbf{i}, \quad -\frac{1}{2}-\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k}, \quad \frac{1}{2}-\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k} \right\} \\
-\mathbf{i}H &= \left\{ -\mathbf{i}, \quad \frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k}, \quad -\frac{1}{2}+\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k} \right\} \\
\mathbf{j}H &= \left\{ \quad \mathbf{j}, \quad -\frac{1}{2}+\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k}, \quad \frac{1}{2}-\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k} \right\} \\
-\mathbf{j}H &= \left\{ -\mathbf{j}, \quad \frac{1}{2}-\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k}, \quad -\frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k} \right\} \\
\mathbf{k}H &= \left\{ \quad \mathbf{k}, \quad -\frac{1}{2}-\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k}, \quad \frac{1}{2}+\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k} \right\} \\
-\mathbf{k}H &= \left\{ -\mathbf{k}, \quad \frac{1}{2}+\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k}, \quad -\frac{1}{2}-\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k} \right\}.
\end{aligned}
\tag{7}
$$

We will see by some general theory that we got them all. For instance, if we look at say

$$\omega H = \left(\frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{k}\right)H$$

we get

$$
\left\{\frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{k}, \left(\frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{k}\right) * \left(-\frac{1}{2}+\frac{1}{2}\overset{x}{\overset{||}{\mathbf{i}}}+\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k}\right),\right.
$$

$$
\left.\left(\frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{k}\right) * \left(-\frac{1}{2}-\frac{1}{2}\overset{x^2}{\overset{||}{\mathbf{i}}}-\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k}\right)\right\}
$$

which involves two calculations which I certainly don't feel like doing directly, so let's use a trick. If, as usual, we let ω denote

$$\frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{k},$$

then

$$-\frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k} = -\overline{\omega}.$$

Therefore, we have

$$\omega(-\overline{\omega}) = -\omega\overline{\omega} = -\|\omega\|^2 = -1$$

which is the first calculation, while the second calculation involves $\omega(-\omega)$. But this is

$$\omega(-\omega) = -\omega^2 = -\left(-\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}\right) = \frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k}.$$

Therefore,

$$\omega H = \left\{\omega,\ -1,\ -\omega^2\right\} = \left\{\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}, -1, \frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k}\right\} = -H.$$

Notice in each example given, the set of cosets partition the group (in the sense that every element of the group is one and only one coset). This is not a coincidence. Of course the easiest way to show that we have a partition is to define an equivalence relation whose equivalence classes are exactly equal to the set of cosets. What could this be? Well, if it were an equivalence relation, then we would have $a \in bH$ if and only a and b were related. However, bH is the set of translates of H by b and so we would have $a = bh$ for some $h \in H$ or $b^{-1}a \in H$. This motivates the following definition:

DEFINITION 10.44. We define a relation on G by $a \equiv_H b$ iff $b^{-1}a \in H$.

The following shows that $\equiv_H$ is an equivalence relation. We denote a class of this equivalence relation by a/H instead of the more cumbersome $a/\equiv_H$. We denote the set of all classes by G/H (instead of the cumbersome $G/\equiv_H$). In more detail, we have that

$$a/_H = \{b \in G : a \equiv_H b\} = \{b \in G : b^{-1}a \in H\} \tag{8}$$

THEOREM 10.45.

(1) *The relation $\equiv_H$ is an equivalence relation.*

(2) *The equivalence classes of this equivalence relation are the left cosets. More precisely, we have that for any $g \in G$ that*

$$gH = g/_H.$$

PROOF.

(1) This is the case because cosets are orbits of the action of H on G given by $h \bullet g = gh^{-1}$. We give a more traditional proof. (zzzzzzzz)

- For any element $g \in G$ we have that $g \equiv_H g$ since $g^{-1}g = e_G \in H$ and the relation is reflexive.
- Suppose that $a \equiv_H b$. Then $b^{-1}a \in H$. But since $H \preccurlyeq G$ we also have that $\left(b^{-1}a\right)^{-1} \in H$. But this is (by socks and shoes) $a^{-1}b$. Hence $a^{-1}b \in H$ and therefore (by definition) $b \equiv_H a$ and the relation is symmetric.
- Suppose that $a \equiv_H b$ and $b \equiv_H c$. Then we have that $b^{-1}a$ and $c^{-1}b$ are in H. Therefore, by closure of the product in H so is $\left(c^{-1}b\right)\left(b^{-1}a\right)$. But this product is $c^{-1}a$. Therefore we have that $a \equiv_H c$ and the relation is transitive.

(2) We are trying to show the equality of two sets, so we need to show mutual containment. We have

$$\begin{aligned} a/H &= \{b \in G : a \equiv_H b\} && \text{(By Equation (8))} \\ &= \{b \in G : b^{-1}a \in H\} \\ &= \{b \in G : b^{-1}a = h \text{ for some } h \in H\} \\ &= \{b \in G : a = bh \text{ for some } h \in H\} \\ &= aH. && \text{(By Equation (6))} \end{aligned}$$

Notice, this means that we have found all of the cosets in Example 10.43**(3)**. ■

Definition 10.46. The number of left cosets, i.e. the number of elements in the set G/H, $\#\left(G/H\right)$, is called *the index* of H in G.

Lagrange's theorem below, Theorem 10.67 back on page 531, gives us the wonderful formula

$$\#\left(G/H\right) = \frac{|G|}{|H|}$$

which "looks right." (Not that that means anything!)

Remark 10.47. We typically do not use any special symbols for the index, we simply denote as $\#\left(G/H\right)$. However, occasionally, for emphasis, we use the notation $[G : H]$ to denote the index of H in G.

Notation 10.48. Since the two sets gH and g/H are equal, we will only use the notation gH. However, we will keep the notation G/H to describe the set of all cosets. That is

$$G/H = \{gH : g \in G\}.$$

Let's translate over the theorem (Theorem 2.82) about equivalence classes to this situation:

Corollary 10.49. *The following are equivalent*

(1) $b^{-1}a \in H$ *i.e.* $a \equiv_H b$*;*

(2) $aH \cap bH \neq \emptyset$*;*

(3) $aH = bH$*;*

(4) $aH \subseteq bH$*.* ■

In practice, the easiest way of showing that two cosets are equal is by the equivalence of **(1)** and **(3)** in this corollary.

$$aH = bH \Leftrightarrow b^{-1}a \in H.$$

EXAMPLE 10.50. Consider the group

$$G := \left\langle (1\,2\,3\overset{x}{\overset{\shortparallel}{4}}5\,6\,7), (1\,2)\overset{y}{\overset{\shortparallel}{(}}3\,6) \right\rangle$$ [II]

and the subgroup $H := \langle x \rangle$ which consists of the the 7 ($7 = |x|$) powers of x:

$$H = \{(1), (1\,2\,3\,4\,5\,6\,7), (1\,3\,5\,7\,2\,4\,6), (1\,4\,7\,3\,6\,2\,5), (1\,5\,2\,6\,3\,7\,4), (1\,6\,4\,2\,7\,5\,3), (1\,7\,6\,5\,4\,3\,2)\}.$$

For practice, we cut up

$$H,\ xyxyH,\ xyH,\ yxyH,\ yx^3yH\ xyx^2yH$$
$$x^2yH,\ x^3yH,\ x^4yH,\ x^5yH,\ x^6yH$$

into equal "piles."

$b^{-1}a$	(1)	$xyxy$	xy	yxy	yx^3y	xyx^2y	x^2y	x^3y	x^4y	x^5y	x^6y
(1)	(1)										
$xyxy$											
xy											
yxy											
yx^3y											
xyx^2y											
x^2y											
x^3y											
x^4y											
x^5y											
x^6y											

{Finish}

EXAMPLE 10.51. Let n be a positive integer. Then the cosets of the subgroup $n\mathbb{Z}$ in $\mathbb{Z}$ are the elements of $\mathbb{Z}_n$. This is because to integers, a and b, are in the same coset of $n\mathbb{Z}$ iff $a - b \in n\mathbb{Z}$, i.e. iff $n \mid (a - b)$. But this last relation was the equivalence relation we originally used to define the elements of $\mathbb{Z}_n$.

[II] This group will be referred to later, in Definition 12.32 coming up on page 676, as Hol ($\mathbb{Z}_7$).

10.5 The action on cosets

Theorem 10.52. *For any subgroup H, the group G preserves the equivalence relation $\equiv_H$. Therefore, we get a well defined action on the set of left cosets of H in G, G/H.*

Proof. Suppose that $a \equiv_H b$. Then by definition, we have that $b^{-1}a \in H$. But this means that

$$(gb)^{-1} = b^{-1}g^{-1} = b^{-1}a \in H.$$

This proves that

$$a \equiv_H b \Rightarrow ga \equiv_H gb.$$

Therefore, the action of G on itself preserves the equivalence relation and as such the action takes equivalence classes to equivalence classes, which are cosets. ■

Let's be a little bit more concrete and summarize:

Theorem 10.53. *For any group G and any subgroup H of G there is an action of G on G/H given by,*

$$g \bullet aH = gaH.$$ ■

You should think of this action as "ignoring the bullet." We will however, continue to put the bullet in if we want to stress the action nature of this!

Proposition 10.54. *This action of G on G/H is transitive.*

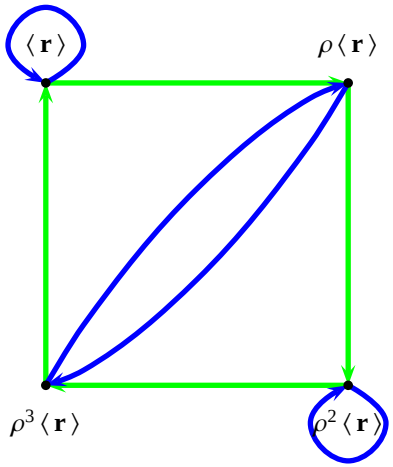
Proof. Every element of G/H has the form gH for some element $g \in G$. But we have that $g \bullet H = gH$ and therefore every coset is in the orbit of H and therefore, the action is transitive. ■

Later we'll see that amazingly, this is, up to the choice of dictionary, the **only** kind of action a group can have. This fact is proven in Theorem 10.92, the orbit stabilizer theorem.

Example 10.55. In Example 10.43**(1)** we saw that

$$\mathcal{D}_4/_{\langle \mathbf{r} \rangle} = \left\{ \langle \mathbf{r} \rangle, \rho \langle \mathbf{r} \rangle, \rho^2 \langle \mathbf{r} \rangle, \rho^3 \langle \mathbf{r} \rangle \right\}.$$

Let's draw the action diagram: (with blue representing the generator $\mathbf{r}$ and green representing the generator ρ)



Where I got the fact that

$$(\mathbf{r}\rho\langle\,\mathbf{r}\,\rangle =)\mathbf{r}\bullet\rho\langle\,\mathbf{r}\,\rangle = \rho^3\langle\,\mathbf{r}\,\rangle$$

since we have that

$$(\rho^3)^{-1}\mathbf{r}\rho = \rho\mathbf{r}\rho = \rho\rho^{-1}\mathbf{r} = \mathbf{r}.$$

Or if you prefer the more direct approach (this approach shouldn't be used once the size of the subgroup gets large!) we have

$$\mathbf{r}\{\rho,\rho\mathbf{r}\} = \{\mathbf{r}\rho,\mathbf{r}\rho\mathbf{r}\} = \{\rho^3\mathbf{r},\rho\}.$$

So if we want to write down the image of generators, we have that ρ goes to $(\,1\,2\,3\,4\,)$ and $\mathbf{r}$ goes to $(\,2\,4\,)$ (after labeling them in a clockwise manner starting at the top-left hand corner). But this is the same action diagram we get when we let $\mathcal{D}_4$ act on the vertices of a square!

This hints at the fact that somehow, the square is "programmed" into the group $\mathcal{D}_4$, even if just abstractly. As a matter of fact, the same sort of thing works in $\mathcal{D}_n$, no matter what n is: take H as the subgroup of $\mathcal{D}_n$ generated by a reflection, ρ. Then the action diagram "looks" the same as the action diagram for the action of $\mathcal{D}_n$ on the regular n–gon.

Recall that by definition, for $H \leqslant G$ and $g \in G$,

$$gHg^{-1} \stackrel{\text{def}}{=} \{ghg^{-1} : h \in H\}.$$

In words this was called the "conjugate of H by g." These gHg^{-1} all turned out to be subgroups of G which are isomorphic to H. The core of $H \leqslant G$, denoted as $H^{\circ G}$, was defined as the intersection of all conjugates.

THEOREM 10.56.

(1) *$gH = H \Leftrightarrow g \in H$ i.e. $G_H = H$.*[I]

(2) *More generally, the stabilizer of gH under the coset action $G \circlearrowright G/H$ is*

$$gHg^{-1}, \qquad i.e \qquad G_{gH} = gHg^{-1}.$$

(3) *The kernel of this action is the core of H in G, $H^{\circ G}$ i.e.*

$$\operatorname{Ker}\left(G \circlearrowright {}^{G}\!/_{H}\right) = H^{\circ G}.$$

(4) *The kernel of the coset action $G \circlearrowright G/H$ is H itself iff H is normal in G.*

(5) *The coset action $G \circlearrowright G/H$ is faithful iff H is* **mal***normal in G.*

PROOF.

(1) For note that since $g \bullet H = gH$ and $H = e_G H$, another way to state that $g \bullet H = H$ is to state that $gH = e_G H$. Then

$$\begin{aligned} gH = e_H H &\Leftrightarrow e_H^{-1} g \in H \\ &\Leftrightarrow g \in H. \end{aligned}$$

(2) By **(1)**, the stabilizer of H under the action $G \circlearrowright G/H$, which in symbols is G_H, is just H itself. Now the rest is simply a consequence of Theorem 10.34.1 indeed:

$$\begin{aligned} \{\overline{g} \in G : \overline{g}gH = gH\} &= G_{g\bullet H} && \text{(Simply what notation means)} \\ &= gG_H g^{-1} && \text{(Theorem 10.34.1)} \\ &= gHg^{-1}. && (G_H = H \text{ by } \mathbf{(1)}) \end{aligned}$$

[I] Or said in another way the stabilizer of $e_G \cdot H$ under the coset action $G \circlearrowright G/H$ is H.

The other parts are an exercise in definitions.

(3) By definition of the kernel of an action (applied in this transitive situation),

$$\begin{aligned}\mathrm{Ker}\left(G \circlearrowright {}^{G}\!/_{H}\right) &= \bigcap_{g\in G} G_{g\bullet H} \\ &= \bigcap_{g\in G} gHg^{-1} \quad \text{(By (2))} \\ &= H^{\circ G}. \quad \text{(Def'n of core)}\end{aligned}$$

(4) By **(3)**, the kernel of the action $G \circlearrowright G/H$ is the core of H in G, $H^{\circ G}$, and by the definition of a normal subgroup, $H^{\circ G} = H$ iff $H \trianglelefteq G$. Hence the kernel of the action is H iff H is normal in G.

(5) This is the same proof as **(4)**, simply replace the words "malnormal" with "normal" and core equal to H to core equal to e_G.

■

Remark 10.57. So if $H \trianglelefteq G$ we should think of this as saying we can "ignore" the contribution of H on G/H. For instance, if we have g_1 and g_2 are arbitrary elements of G and h in some normal subgroup H we would have that each of hg_1g_2H, g_1hg_2H, and g_1g_2hH are simply g_1g_2H. If however H were not normal, we would only be able to ignore elements of H on the right. For instance, in this situation, of the 3 cosets listed, only g_1g_2hH would equal g_1g_2H in general.

Exercise 10.32.

Solution

(1) Show that if $N \trianglelefteq G$, then

$$g \bullet N = \overline{g} \bullet N \implies \forall a \in G \quad g \bullet aN = \overline{g} \bullet aN.$$

(This says that $\rho^{\bullet}(g)$ is determined by what it does to a single coset.)

(2) Show that this doesn't hold in Example **??**. Why doesn't this contradict the first part of this exercise?

Remark 10.58. We therefore have that for any malnormal subgroup H of index n in G we have a subgroup of $\mathcal{S}_n$ which is isomorphic to H. Namely, the image of H under the injective homomorphism

$$\rho^{\bullet} : G \to \mathcal{S}_{G/H} \cong \mathcal{S}_n.$$

The orbit stabilizer theorem, proven later, will show that the converse is also true.

Corollary 10.59. *A subgroup is normal iff it is the kernel of some homomorphism* (*namely the permutation representation of the action*).

Proof. The kernel of any homomorphism is always normal (Theorem 8.49). The converse follows from the theorem. ■

Theorem 10.60. *Given $H \leqslant G$ we have that H is*

(1) *Normal in G iff for every $C \in G/H$ we have for every $h \in H$, $h \bullet C = C$.*

(2) *Malnormal in G iff $G \circlearrowleft G/H$ is faithful.*

Proof. These are just restatements of Theorem 10.56. ■

Theorem 10.61. *An index two subgroup of a group G is normal in G.* (*That is if the order of a subgroup is half the order of a group.*)

Proof. We show that in this case, H is the kernel of the action of $G \circlearrowleft G/H$. We label the two cosets in G/H by $\mathrm{I} = H$ and II for the coset not in H. Then, for any $h \in H$ we have that $h \bullet \mathrm{I} = \mathrm{I}$. But since $h\bullet$ is a bijection, we have to have that $h \bullet \mathrm{II} = \mathrm{II}$ as well. ■

For $N \trianglelefteq G$ there is a method that makes calculating when two cosets are equal a bit easier. It actually cuts down the number of calculations one needs to do in half. In several books it is (the basis for) the definition of a normal subgroup.

Proposition 10.62. *Suppose that $N \trianglelefteq G$. Then $aN = bN$ holds iff $ab^{-1} \in N$.*

Remark 10.63. Since $aN = bN$ holds iff $a \equiv_n b$, i.e. $b^{-1}a \in N$, this proposition could be thought of as saying that

$$ab^{-1} \in N \Leftrightarrow b^{-1}a \in N.$$

Remark 10.64. The converse of the proposition is also true. Therefore its easy to find counterexamples when N is not normal. For instance, let's take $G = \mathcal{S}_3$ and consider the non-normal subgroup $\langle (1\,2) \rangle = \{(1), (1\,2)\}$. Then

$$(1\,2\,3)^{-1}(1\,3) = (1\,3\,2)(1\,3) = (1\,3)(2\,3)(1\,3) = (2\,1) = (1\,2) \in \langle (1\,2) \rangle$$

but

$$(1\,3)(1\,2\,3)^{-1} = (1\,3)(1\,3\,2) = (1\,3)(1\,3)(2\,3) = (2\,3) \notin \langle (1\,2) \rangle.$$

Proof. Since N is normal by assumption and $a \in G$ we have that if $ab^{-1} \in N$, then

$$b^{-1}a = a^{-1}\left(ab^{-1}\right)a \in N.$$

If on the other hand we had that $b^{-1}a \in N$ then conjugating by a and using normality of N would show that ab^{-1} is also in N. ■

In this exercise we discuss right cosets which are typically discussed a lot in other texts but are downplayed in this text.

Exercise 10.33.

(1) Show that the relation on G defined by $a \ {}_H\!\equiv b$ iff $ab^{-1} \in H$ is an equivalence relation (this is what we have dealt with in the previous example).

SOLUTION

(2) Show that the equivalence class of $g \in G$ under this relation is

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

called a *right coset* of H in G .

(3) Show that $H \trianglelefteq G$ iff $gH = Hg$ for every $g \in G$.
Hint! Let me start part of this for you. Suppose that $H \trianglelefteq G$ and let $k \in gH$. Then $k = gh$ for some $h \in H$. But since H is normal in G, we know that $ghg^{-1} = \overline{h} \in H$. As such

$$k = gh = ghg^{-1}g = \overline{h}g \in Hg.$$

EXERCISE 10.34. Prove TFAE

SOLUTION

(1) $gH = Hg$
(2) gH is stabilized by H in the coset action, that is $\forall h \in H, h \bullet gH = gH$[II]
(3) $g \in \mathrm{N}_G(H)$ (this was defined in Exercise 8.40 back on page 411.).

REMARK 10.65. We have built up a host of equivalent ways of showing that a subgroup N is normal in G. They are

(1) For all $g \in G$ and $h \in N$, $ghg^{-1} \in N$.

(2) The core of N, is equal to N. This was the original definition which in symbols can be written as

$$N^{\circ G} = N.$$

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

(4) For all $g \in G$, $gN = Ng$.[III]

(5) There exists some homomorphism with domain G and kernel N.

(6) The coset action of N on G/N is trivial. This means for all $h \in N$ and $g \in G$ we have that $h \bullet gN = gN$, i.e. $hgN = gN$.

(7) The coset action of N on G/N fixes every non-identity coset. This means for all $h \in N$ and $g \notin H$ we have that $h \bullet gN = gN$, i.e. $hgN = gN$. This is clearly equivalent to the previous since we always get that N fixes the coset $e_G N$ independent of whether or not N is normal in G.

(8) The normalizer of N, defined in Exercise 8.40, is equal to all of G. In symbols, $\mathrm{N}_G(N) = G$.

We will see in the quotient groups chapter one more very important (and the way that normal subgroups are applied in practice!) way we could have defined what it means for a subgroup to be normal! It turns out that N is normal in G iff we can define a group structure on G/N (in a natural way).

[II] Said in another way, $H \subseteq G_{gH}$

10.6 Subgroups of $\mathbb{Z}$ (reprieve)

One case where we can classify all subgroups of a given group is the case when the group is taken as the integers (under addition of course).

We have already proven that every subgroup of $\mathbb{Z}$ is of the form $n\mathbb{Z}$ in 6.22. For the purposes of given an application of the fairly abstract theorems, we give a very different proof.

Theorem 10.66. *Let $H \leqslant \mathbb{Z}$. Then $H = n\mathbb{Z}$ for some integer n.*

Proof. Since $\mathbb{Z}$ is abelian, H is a normal subgroup of $\mathbb{Z}$. By Theorem 10.59 there is some group G and some homomorphism

$$\varphi : \mathbb{Z} \to G$$

whose kernel is H, i.e. $\operatorname{Ker}(\varphi) = H$. Let $g = \varphi(1)$ and let $n = |g|$ (if n is infinite, the kernel, and hence H, is trivial). Then $\varphi(m) = g^m$. Thus $m \in \operatorname{Ker}(\varphi)$ iff $g^m = e_G$. But this happens iff n divides m, i.e. iff $m \in n\mathbb{Z}$. Hence $H = \operatorname{Ker}(\varphi) = n\mathbb{Z}$ as desired. ∎

10.7 Lagrange's Theorem

Theorem 10.67 (Lagrange's Theorem)**.** *Suppose that G is a finite group and H is any subgroup of G. Then*

$$\frac{|G|}{|H|} = \#\left(G/H\right).$$

In particular, since $\#(G/H)$ is an integer, we have that the order of H divides the order of G.

Proof. G acts transitively on the set set of left cosets, G/H. Thus all left cosets have the same number of elements. Since H is a left coset, $H = e_G H$, all left cosets have $|H|$ elements. Thus $|G| = \#(G/H) \cdot |H|$ as desired. ∎

Remark 10.68. The converse of Lagrange's theorem is **FALSE**. For instance 6 divides the order of $\mathcal{A}_4$ (which is 12) but $\mathcal{A}_4$ has no subgroup of order 6 (index 2). For $\mathcal{A}_5$ we will see that there are numerous counter examples. For instance, $\mathcal{A}_5$ has no subgroups of orders 30, 20, or 15. You should think of Lagrange's theorem as narrowing down the possible orders of a potential subgroup of a group.

Exercise 10.35. **T** or **F**:

(1) Lagrange's implies a group of order 72 has a subgroup of order 24.

(2) Lagrange's implies a group of order 72 doesn't have a subgroup of order 24.

(3) Lagrange's implies a group of order 72 doesn't have a subgroup of order 16.

(4) Lagrange's implies a group of order 72 has an element of order 16.

Solution

Corollary 10.69. *For any finite group G,*

(1) *If $g \in G$, then $|g| \Big| |G|$.*

(2) *If G is a group of order n, then for any element $g \in G$, $g^n = e_G$.*

(3) *If G is a group whose order is equal to a prime number p, then G is cyclic of order p.*

(4) *Let p be a prime number. Up to isomorphism, there is one and only one group of order p, namely $\mathbb{Z}_p$.*

Proof.

(1) Let $g \in G$. Then $\langle g \rangle \leqslant G$. Hence, by Lagrange's theorem like every subgroup of G, $|\langle g \rangle|$ divides $|G|$. But by Corollary 5.61 we know that $|\langle g \rangle| = |g|$. Therefore, substituting, $|g|$ divides $|G|$.

(2) Let $g \in G$. By the first part, we have that $n = |g|a$ for some integer a. Therefore

$$\begin{aligned} g^n &= g^{|g|a} \\ &= \left(g^{|g|}\right)^a \\ &= e_G^a \\ &= e_G. \end{aligned}$$

(3) Since 1 is not a prime number, there exists some element $g \in G$ which does not equal the identity. Since $|g| \Big| |G| = p$, and $|g| \neq 1$, we must have that $|g| = p$. Since $|g| = |\langle g \rangle|$ (again by Corollary 5.61) we must have that $\langle g \rangle = G$. As such, G is generated by g and G is cyclic.

(4) By **(??)** we know that any group of order p is cyclic of order p. We have shown in Corollary 8.61 back on page 419 that every cyclic group of order p is isomorphic to $\mathbb{Z}_p$, therefore every group of order p, being cyclic, is isomorphic to $\mathbb{Z}_p$. ∎

Remark 10.70. This means no matter what group you come across of order p that group must be isomorphic to $\mathbb{Z}_p$. So if you see a group, say, with 13 elements in it, you know without giving it anymore thought that that group is isomorphic to $\mathbb{Z}_{13}$.

Exercise 10.36. Suppose that $H, K \leqslant G$ have relatively prime finite orders. Prove that $H \cap K = \{e_G\}$.

Solution

Exercise 10.37. Suppose that $H, K \leqslant G$ (all finite) and $H \cap K = \{e_G\}$.

Solution

(1) Prove that

$$\{kH : k \in K\} \tag{9}$$

make up distinct cosets of H in G.

(2) In addition, assume that $|G| \leq |H| \cdot |K|$. Prove that the set of cosets in (9) is equal to G/H.

Exercise 10.38. Use Exercise 10.37 to prove that if $H, K \leqslant G$ with $|H| \cdot |K| > |G|$, then $H \cap K$ contains a non-identity element.

SOLUTION

Exercise 10.39. Suppose that $P \leqslant G$ has prime order, and that $H \leqslant G$. Prove that if

$$H \cap P \neq \{e_G\} \Rightarrow P \subseteq H.$$

SOLUTION

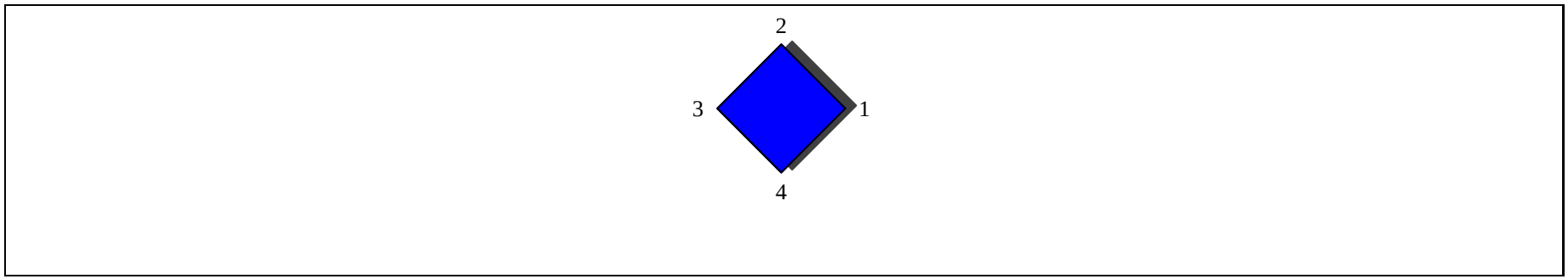
Exercise 10.40. Let G be a finite group. Suppose that there is a $g \in G$ so that

- $|g|$ is prime,
- g is self centralizing. That is $C_G(g) = \langle g \rangle$,
- g is not-central. That is $\exists h \in G$ s.t. g and h do not commute.

Then prove that G is centerless.
Hint! Use the previous exercise with $P = C_G(g)$ and $H = C_G(h)$.

SOLUTION

Example 10.71. We give an application of Exercise 10.37. Consider $\widetilde{\mathcal{D}_4} := \langle (1\,2\,3\,4), (2\,4) \rangle \leqslant \mathcal{S}_4$ which get by identifying the symmetries of a square, $\mathcal{D}_4$, with the permutations of its (the square's) vertices labeled counterclockwise starting at 1. (More formally, we are using the permutation representation of the faithful action of $\mathcal{D}_4 \curvearrowright \{\{4\}\}$.) Since $\mathcal{D}_4$ does not contain any elements of order 4, the group $\langle (1\,2\,3) \rangle$ only intersects $\widetilde{\mathcal{D}_4}$ trivially. (Alternatively, this is because 8 and 3 are relatively prime).



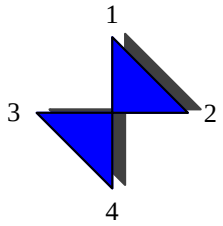
Therefore the three cosets of $\widetilde{\mathcal{D}_4}$ in $\mathcal{S}_4$ can be represented as

$$\mathcal{S}_4/\widetilde{\mathcal{D}_4} = \{\widetilde{\mathcal{D}_4},\ (1\,2\,3)\widetilde{\mathcal{D}_4},\ (1\,3\,2)\widetilde{\mathcal{D}_4}\}.$$

Since 24 isn't that large of a number, we could take the time (probably 5 or 10 minutes) to write out all 8 elements of $\widetilde{\mathcal{D}_4}$ and then figure, then take these 8 elements and multiply on the left by $(1\,2\,3)$ to figure out the 8 elements of the second coset listed, $(1\,2\,3)\widetilde{\mathcal{D}_4}$, and then either multiply these 8 elements again by $(1\,2\,3)$ or multiply the

original 8 elements in $\widetilde{\mathcal{D}_4}$ by $(1\ 3\ 2)$ we could figure out the 8 elements in the third coset listed $(1\ 3\ 2)\,\widetilde{\mathcal{D}_4}$. But for some practice, let's attempt to do this most abstractly in a few special cases.

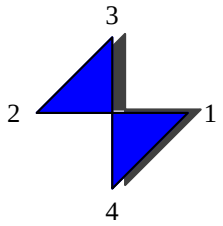
- Which coset does the 2–cycle $(1\ 2)$ belong? Well since $(1\ 2)$ breaks the square



it is not in $\widetilde{\mathcal{D}_4}$.
Is it in $(1\ 2\ 3)\,\widetilde{\mathcal{D}_4}$? Well, we know that it is iff $(1\ 2\ 3)^{-1}(1\ 2) \in \widetilde{\mathcal{D}_4}$. But

$$(1\ 2\ 3)^{-1}(1\ 2) = (2\ 3)(1\ 2)(1\ 2) = (2\ 3),$$

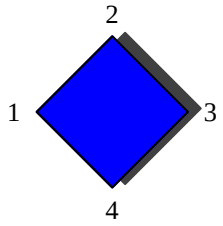
and again, this is not a symmetry.



Therefore, since we know that every element of $\mathcal{S}_4$ is in some (actually a unique) coset, $(1\ 2) \in (1\ 3\ 2)\,\widetilde{\mathcal{D}_4}$ by a process of elimination. For some more practice, let's verify this. Indeed, since

$$(1\ 3\ 2)^{-1}(1\ 2) = (1\ 2)(2\ 3)(1\ 2) = (1\ 3),$$

and since $(1\ 3)$ does indeed correspond to a symmetry of the square:



we have proven directly that

$$(1\,2) \in (1\,3\,2)\widetilde{\mathcal{D}_4} \quad \text{and} \quad (1\,2)\widetilde{\mathcal{D}_4} = (1\,3\,2)\widetilde{\mathcal{D}_4}. \tag{10}$$

REMARK 10.72. Of course I did all of this extra work for your benifit! If on a homework problem you were asked "what coset is $(1\,2)$ in" you would simply state:

$(1\,2) \in (1\,3\,2)\widetilde{\mathcal{D}_4}$ since $(1\,3\,2)^{-1}(1\,2) = (1\,3) \in \widetilde{\mathcal{D}_4}$.

You **would not** bother to write down on your official homework all of the steps that did not work! In fact, in many problems you should be able to "guess" what will work just from the problems context. For instance, if a problem asks you to prove to groups are isomorphic, then don't waste your time trying possibilities which would end up proving that the groups are not isomorphic!

- How about the 3–cycle $(2\,3\,4)$? We could go through the same process as before, but let's use a couple of tricks this time around. We know that $(1\,2\,3\,4) \in \widetilde{\mathcal{D}_4}$. Indeed, it was one of the given generators of this group. Therefore, using the properties of the action $\mathcal{S}_4 \circlearrowright \mathcal{S}_4/\widetilde{\mathcal{D}_4}$ we have

$$\begin{aligned}
(1\,2\,3\,4)\widetilde{\mathcal{D}_4} = \widetilde{\mathcal{D}_4} &\Rightarrow (1\,2)(2\,3\,4)\widetilde{\mathcal{D}_4} = \widetilde{\mathcal{D}_4} && ((1\,2\,3\,4) = (1\,2)(2\,3\,4)) \\
&\Rightarrow (2\,3\,4)\widetilde{\mathcal{D}_4} = (1\,2)\widetilde{\mathcal{D}_4} && \text{(Hit bs on left by } (1\,2)) \\
&\Rightarrow (2\,3\,4)\widetilde{\mathcal{D}_4} = (1\,3\,2)\widetilde{\mathcal{D}_4}. && \text{(Equation (10))}
\end{aligned}$$

For an application of the previous exercise let's reprove that $\mathcal{S}_4$ and $\mathcal{A}_4$ are centerless. Indeed, any 3–cycle is self-centralizing and non-central and both of these groups contain the 3–cycle $(1\,2\,3)$. Another application could be to prove that $\mathbf{SL}_3(\mathbb{Z}_2)$ is also centerless by proving that only powers of the matrix (which has order 3)

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

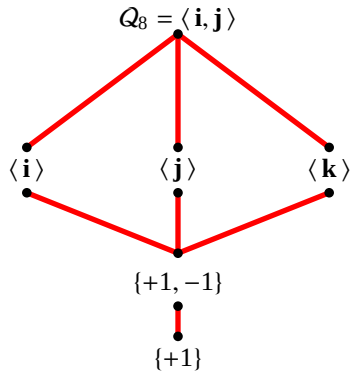
commute with it.

EXAMPLE 10.73. Since $\mathcal{D}_6$ has order $6 \cdot 2 = 12$, the only possibly subgroups of $\mathcal{D}_6$ have order 1, 2, 3, 4, 6 or 12.

EXAMPLE 10.74. Let's find all proper subgroups of Q_8. By Lagrange's theorem, such a subgroup has order 2 or 4. If it has order two it must be cyclic and generated by an element of order 2. There is only one such such element, namely -1. Now all other non-identity elements of Q_8 have order 4. Hence the group generated by any such element has order 4. Thus there are only 4 proper subgroups of Q_8 and they are

$$\{1,-1\} = \langle -1 \rangle, \{1,-1,\mathbf{i},-\mathbf{i}\} = \langle \mathbf{i} \rangle, \{1,-1,\mathbf{j},-\mathbf{j}\} = \langle \mathbf{j} \rangle, \{1,-1,\mathbf{k},-\mathbf{k}\} = \langle \mathbf{k} \rangle$$

From this we can make what is known as a subgroup lattice diagram, which is simply the poset diagram of the set of subgroups. For instance for Q_8 we have



EXERCISE 10.41.

SOLUTION

(1) Suppose that H_1 and H_2 are subgroups of G of relatively prime order. Show that

$$H_1 \cap H_2 = \{e_G\}.$$

(2) Show that the least common multiple of the order of H_1 and H_2 must divide the order of $\langle H_1, H_2 \rangle$. Find an example were this is exactly true and an example where it is not exactly true.

(3) Suppose that H has prime order and is a subgroup of G. Show that H is either normal or malnormal.

(4) Find all subgroups of $B\mathcal{D}_3$, $\mathcal{A}_4$ and $\mathcal{S}_5$ and make their subgroup lattice diagrams (for $\mathcal{A}_4$, you want to keep in mind that there are no subgroups of order 6!)

10.8 BACK TO PUSHING AROUND COSETS

EXAMPLE 10.75. Consider again the binary tetrahedra group and the subgroup H:

$$G = B\left(\blacktriangleright^3\right) \quad \text{and} \quad H = \left\langle -\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k} \right\rangle.$$

In Example 10.43**(??)**, we saw that there are 8 cosets of H in G (i.e. the elements of G/H) which can conveniently be labeled as

$$1 = H \qquad 2 = (-1)H \qquad 3 = \mathbf{i}H \qquad 4 = -\mathbf{i}H \tag{11}$$
$$5 = \mathbf{j}H \qquad 6 = -\mathbf{j}H \qquad 7 = \mathbf{k}H \qquad 8 = -\mathbf{k}H. \tag{12}$$

We typically think of $B\left(\blacktriangleright^3\right)$ as being generated as:

$$\left\{\frac{1}{2}+\frac{1}{2}\mathbf{i}+\overset{\omega}{\overset{\shortparallel}{\frac{1}{2}}}\mathbf{i}+\frac{1}{2}\mathbf{k},\ \mathbf{i}\right\}.$$

However, it is slightly more convenient to take our generating set as the set $\{\mathbf{i}, x\}$ instead (since $\omega = \mathbf{i}^2x^2 = -x^2$ this is indeed a generating set). It's easy to figure out how $\mathbf{i}$ acts on these cosets since it is exactly how $\mathbf{i}$ hits Q_8 under the Cayley action (since the cosets are labeled by elements of Q_8!). Thus we have

$$\rho^{\bullet}(\mathbf{i}) = (1\,3\,2\,4)(5\,7\,6\,8).$$

Now there are a lot more calculations involved in figuring out $x\bullet$ however. Let's start with the easiest 2:

$$\begin{aligned} x \bullet H &= xH = H && \text{(Since } x \in H\text{)}\\ x \bullet (-1)H &= (-1)xH = (-1)H. && \text{(Since } x \in H\text{)} \end{aligned}$$

Now to figure out what goes on with everybody else, let's do a couple of preliminary calculations:

$$\begin{aligned} \mathbf{i}^{-1}x\mathbf{i} &= \mathbf{i}x\mathbf{i}^{-1}\\ &= \mathbf{i}\left(-\frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k}\right)\mathbf{i}^{-1}\\ &= \mathbf{i}\left(-\frac{1}{2}\right)\mathbf{i}^{-1}+\mathbf{i}\left(\frac{\mathbf{i}}{2}\right)\mathbf{i}^{-1}+\mathbf{i}\left(\frac{\mathbf{j}}{2}\right)\mathbf{i}^{-1}+\mathbf{i}\left(\frac{\mathbf{k}}{2}\right)\mathbf{i}^{-1}\\ &= -\frac{1}{2}+\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k}. \end{aligned}$$

and similarly

$$\begin{aligned} \mathbf{j}^{-1}x\mathbf{j} = \mathbf{j}x\mathbf{j}^{-1} &= -\frac{1}{2}-\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k}\\ \mathbf{k}^{-1}x\mathbf{k} = \mathbf{k}x\mathbf{k}^{-1} &= -\frac{1}{2}-\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k}. \end{aligned}$$

Next, if we look back to the table of cosets for H in G, (7), we see that

$$-\frac{1}{2}+\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k} \in \mathbf{j}H \Rightarrow \left(-\frac{1}{2}+\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k}\right)H = \mathbf{j}H \tag{13}$$

$$-\frac{1}{2}-\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k} \in \mathbf{k}H \Rightarrow \left(-\frac{1}{2}-\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{j}-\frac{1}{2}\mathbf{k}\right)H = \mathbf{k}H \tag{14}$$

$$-\frac{1}{2}-\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k} \in \mathbf{i}H \Rightarrow \left(-\frac{1}{2}-\frac{1}{2}\mathbf{i}-\frac{1}{2}\mathbf{j}+\frac{1}{2}\mathbf{k}\right)H = \mathbf{i}H. \tag{15}$$

These calculations are of considerable help since now we can now see

$$\begin{aligned} x \bullet \mathbf{i}H &= x\mathbf{i}\,H \\ &= \mathbf{i}\mathbf{i}^{-1}x\mathbf{i}\,H \\ &= \mathbf{i} \bullet \left(\mathbf{i}^{-1}x\mathbf{i}\right)H \\ &= \mathbf{i} \bullet \left(-\frac{1}{2} + \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k}\right)H \\ &= \mathbf{i} \bullet (\mathbf{j}\,H) \qquad \text{(By (13))} \\ &= \mathbf{k}\,H \end{aligned}$$

and

$$\begin{aligned} x \bullet \mathbf{k}H &= x\mathbf{k}\,H \\ &= \mathbf{k}\mathbf{k}^{-1}x\mathbf{k}\,H \\ &= \mathbf{k} \bullet (\mathbf{k}^{-1}x\mathbf{k}) \bullet H \\ &= \mathbf{k} \bullet \left(-\frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}\right)H \\ &= \mathbf{k} \bullet (\mathbf{i}\,H) \qquad \text{(By (14))} \\ &= \mathbf{k}\mathbf{i}\,H. \\ &= \mathbf{j}\,H. \end{aligned}$$

and lastly, since x has order 3 we can save ourselves a calculation and conclude $x \bullet \mathbf{j}\,H = \mathbf{i}\,H$. Sticking negatives in we can also see that

$$x^3 \bullet (-\mathbf{i}H) = x^2 \bullet (-\mathbf{k}H) = x \bullet (-\mathbf{j}H) = -iH.$$

Going to our dictionary in Equation (11) we have

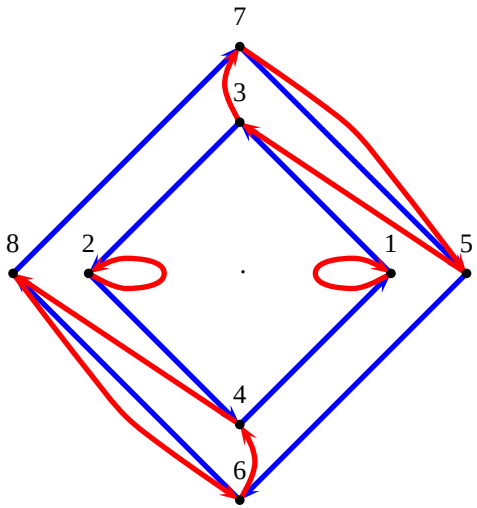
$$\rho^\bullet(x) = (1)(2)(3\,7\,5)(4\,8\,6) = (3\,7\,5)(4\,8\,6).$$

Finally, since $\mathbf{i}\,H\mathbf{i}^{-1} \cap H = \{1\}$, we have that $H^{\circ G} = \{1\}$ and so H is malnormal. So we have that $\rho^\bullet$ is injective and

$$B\left(\blacktriangleright^3\right) \cong \mathrm{Im}\,(\rho^\bullet) = \left\langle (1\,3\,2\,4)\overset{\mathbf{i}}{\downarrow}(5\,7\,6\,8), (3\,7\,5)\overset{x}{\downarrow}(4\,8\,6) \right\rangle.$$ [I]

Let's draw the action diagram (blue=**i** and red=x):

[I] One cool thing is that having turned this group into a symmetric group, we can now enter it into `MAPLE` (since `MAPLE` understands only symmetric groups!).



EXERCISE 10.42. Show that

$$\mathbf{J} = \begin{pmatrix} [0]_3 & -[1]_3 \\ [1]_3 & [0]_3 \end{pmatrix} \quad \text{and} \quad \mathfrak{Q} = \begin{pmatrix} [1]_3 & [1]_3 \\ [0]_3 & [1]_3 \end{pmatrix}$$

have orders 4 and 3 and generate $\mathbf{SL}_2(\mathbb{Z}_3)$.

SOLUTION

EXAMPLE 10.76. Let's use the previous example to show that $B(\blacktriangleright^3)$ is isomorphic to $\mathbf{SL}_2(\mathbb{Z}_3)$ (the notation is as in the previous exercise). This latter group acts faithfully on the set of $3^2 - 1 = 8$ non-zero column vectors. We want to match what we did in the previous example. So we better use the generators that have orders 4 and 3.

Now the matrix $\mathbf{J}$ cuts up the 8 elements of

$$\mathbb{Z}_3^2 \setminus \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$$

into two cycles as follows:

$$\mathbf{J}^4 \bullet \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \mathbf{J}^3 \bullet \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \mathbf{J}^2 \bullet \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \mathbf{J} \bullet \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and

$$\mathbf{J}^4 \bullet \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \mathbf{J}^3 \bullet \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \mathbf{J}^2 \bullet \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \mathbf{J} \bullet \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Meanwhile, $\mathfrak{Q}$ fixes both of the vectors

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

and has two other (non-trivial) cycles

$$\mathfrak{Q}^3 \bullet \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \mathfrak{Q}^2 \bullet \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathfrak{Q} \bullet \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and

$$\mathfrak{Q}^3 \bullet \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \mathfrak{Q}^2 \bullet \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \mathfrak{Q} \bullet \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

Now we want to be able to draw the action diagram in such a way that its obvious that they line up. Of course I could just write it down and you can visually see that I did it correctly, let's take a few moments to explain how to actually do this in an efficient manner. This can be trickier than it sounds, so we proceed in a sensible way. How should we form a dictionary on

$$\mathbb{Z}_3^2 \setminus \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

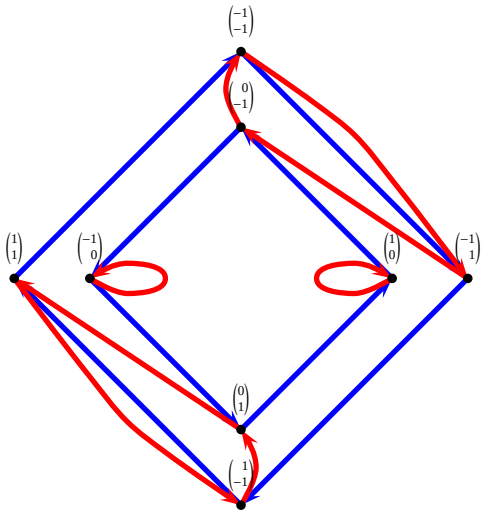
so that the action diagrams match up (given the dictionary we already formed on the cosets for $B\left(\blacktriangleright^3\right)$)?

Since we labeled the two fixed cosets of our generator of order 3, ω^2 (which were H and $-H$) as 1 and 2 respectively, we need to line 1 and 2 up with the two fixed points of our element of order 3, $\mathfrak{Q}$, in $\mathbf{SL}_2(\mathbb{Z}_3)$:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

So, let's say we line up 1 with the first vector and 2 with the second. Now following the orbit of **J** around, we will need to label $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$ as 3. In our first indication we are doing things correctly, notice that $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$ lines up with 2, which luckily is what we lined up with 2 in the first place. The last bit of information we can get from **J** for now is that we need to put $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ as 4.

Now in the original diagram, the element of order 3 took the coset labeled 3 to the coset labeled 7. So therefore, we label what $\mathfrak{Q}$ does to what we labeled as 3 here as 7 in our new situation. This is the vector $\begin{bmatrix} -1 \\ -1 \end{bmatrix}$ Now we use **J** to push this vector around while matching it up with the previous situation. After finishing all this, we still need to check that things line up properly, this is where cross our fingers and hope that everything we have just done is not a waste of time. Let's draw out the entire action diagram now:



Note that we got lucky since the two diagrams, the one for $B\left(\blacktriangleright^3\right)$ and the one right above, match exactly (if we replace the vectors with the appropriate numbers)! Hence if we make the dictionary official, then we have just proven that

$$\mathbf{SL}_2(\mathbb{Z}_3) \cong \left\langle \overset{\mathbf{J}}{\downarrow}(1\,3\,2\,4)(5\,7\,6\,8), \overset{\mathbf{c}}{\downarrow}(3\,7\,5)(4\,8\,6) \right\rangle$$

as well. In particular, since both groups are isomorphic to this same group symmetric group, we have (since $\cong$ is an equivalence relation) we have proven that

$$B\left(\blacktriangleright^3\right) \cong \mathbf{SL}_2(\mathbb{Z}_3).$$

In fact we showed that (on generators) the isomorphism is via

$$\mathbf{i} \mapsto \begin{pmatrix} [0]_3 & -[1]_3 \\ [1]_3 & [0]_3 \end{pmatrix} \quad \text{and} \quad -\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k} \mapsto \begin{pmatrix} [1]_3 & [1]_3 \\ [0]_3 & [1]_3 \end{pmatrix}.$$

Remark 10.77. Since $x^{-1}\mathbf{i}x = \mathbf{j}$ we have also found a copy of Q_8 chilling inside of $\mathbf{SL}_2(\mathbb{Z}_3)$, which is generated by

$$\begin{pmatrix} [0]_3 & -[1]_3 \\ [1]_3 & [0]_3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} [1]_3 & -[1]_3 \\ [0]_3 & [1]_3 \end{pmatrix} \cdot \begin{pmatrix} [0]_3 & -[1]_3 \\ [1]_3 & [0]_3 \end{pmatrix} \cdot \begin{pmatrix} [1]_3 & [1]_3 \\ [0]_3 & [1]_3 \end{pmatrix} = \begin{pmatrix} -[1]_3 & [1]_3 \\ [1]_3 & [1]_3 \end{pmatrix}.$$

Example 10.78. Let's deal with a bigger example, namely an example of $\mathcal{S}_5$ acting faithfully and transitively on 6 things, and thus giving an injection

$$\mathcal{S}_5 \to \mathcal{S}_6$$

which will not be the standard one (since $\mathcal{S}_5 \subseteq \mathcal{S}_6$ fixes 6, and so is not a transitive action). In $\mathcal{S}_5$, by Exercise 7.1 back on page 347, there are a total of

$$\frac{5!}{5} = 4! = 24$$

5–cycles. These 24 5–cycles are split up into 6 cyclic subgroups of order 5 (a group of order 5 contains 4 elements of order 5 each) sitting inside of $\mathcal{S}_5$. They are

$$\begin{aligned}
H_1 &:= \langle (1\,2\,3\,4\,5) \rangle \\
H_2 &:= \langle (2\,1\,3\,4\,5) \rangle \\
H_3 &:= \langle (3\,1\,2\,4\,5) \rangle \\
H_4 &:= \langle (3\,2\,1\,4\,5) \rangle \\
H_5 &:= \langle (2\,3\,1\,4\,5) \rangle \\
H_6 &:= \langle (1\,3\,2\,4\,5) \rangle .
\end{aligned}$$

To figure out which of the subgroups a given 5–cycle lives in, we simply raise the given 5–cycle to a power so that 4 and 5 become adjacent and compare the resulting (still a 5-cycle) permutation with the listed generaters of the H_i (each were picked to have 4 and 5 adjacent to each other). For every 5–cycle it's easy to see that there is a unique way of doing this. For example,

$$(1\,3\,4\,2\,5)^2 = (1\,4\,5\,3\,2) = (3\,2\,1\,4\,5) \Rightarrow (1\,3\,4\,2\,5) \in \langle (3\,2\,1\,4\,5) \rangle = H_4.$$

Since the conjugate of a 5–cycle is a 5–cycle, we have that $\mathcal{S}_5$ acts on these 6 subgroups.

Now $\mathcal{S}_5$ is generated by $(1\,2)$ and $(1\,2\,3\,4\,5)$, so we simply need to see how these two elements act. For starters,

$$\begin{aligned}
(1\,2)[(1\,2\,3\,4\,5)](1\,2)^{-1} &= (2\,1\,3\,4\,5) \in H_2 \\
(1\,2)[(3\,1\,2\,4\,5)](1\,2)^{-1} &= (3\,2\,1\,4\,5) \in H_4 \\
(1\,2)[(2\,3\,1\,4\,5)](1\,2)^{-1} &= (1\,3\,2\,4\,5) \in H_6.
\end{aligned}$$

(In this calculation I've used the fact that $(1\,2)$ has order 2 to save a bit of time.) Hence under the action, $(1\,2)$ acts like the 2 2 2–cycle

$$(H_1\ H_2)(H_3\ H_4)(H_5\ H_6).$$ [II]

Now we find the action of the other generator, $(1\,2\,3\,4\,5)$ on the H's.

$$\begin{aligned}
(1\,2\,3\,4\,5)[(1\,2\,3\,4\,5)](1\,2\,3\,4\,5)^{-1} &= (1\,2\,3\,4\,5) \in H_1 && \Rightarrow H_1 \mapsto H_1 \\
(1\,2\,3\,4\,5)[(2\,1\,3\,4\,5)](1\,2\,3\,4\,5)^{-1} &= (3\,2\,4\,5\,1) = (1\,3\,2\,4\,5) \in H_6 && \Rightarrow H_2 \mapsto H_6 \\
(1\,2\,3\,4\,5)[(3\,1\,2\,4\,5)](1\,2\,3\,4\,5)^{-1} &= (4\,2\,3\,5\,1), (4\,2\,3\,5\,1)^3 = (2\,1\,3\,4\,5) \in H_2 && \Rightarrow H_3 \mapsto H_2 \\
(1\,2\,3\,4\,5)[(3\,2\,1\,4\,5)](1\,2\,3\,4\,5)^{-1} &= (4\,3\,2\,5\,1), (4\,2\,3\,5\,1)^3 = (3\,1\,2\,4\,5) \in H_3 && \Rightarrow H_4 \mapsto H_3
\end{aligned}$$

and since $(1\,2\,3\,4\,5)$ has order 5, we can save a bit of time and use general theory to conclude that it sends $H_5 \mapsto H_4$ and $H_6 \mapsto H_5$. Hence $(1\,2\,3\,4\,5)$ acts as the permutation

$$(H_1)(H_2\ H_6\ H_5\ H_4\ H_3).$$

Therefore, using the implied dictionary from reading the indices of the subgroups, we have a homomorphism

$$\begin{aligned}
\varphi &: \mathcal{S}_5 \to \mathcal{S}_6 \\
&: (1\,2) \mapsto (1\,2)(3\,4)(5\,6) \\
&: (1\,2\,3\,4\,5) \mapsto (2\,6\,5\,4\,3).
\end{aligned}$$

[II] Of course this isn't a coincidence. When I choose how to order the H's I fudged this so it would work out!

Now we still need to show that the homomorphism is an injection, i.e. that the action is faithful. We leave this to you in Exercise 10.44 where you show that all (non-trivial) cycle types in $\mathcal{S}_5$ go to a non-trivial cycle type in $\mathcal{S}_6$.

Alternatively, we cheat and use a future result which states that the only PNT, normal subgroup of $\mathcal{S}_5$ is $\mathcal{A}_5$. Therefore, if the action were not faithful, since the kernel of any homomorphism is normal, it would have to contain each and every even permutation, i.e. all of $\mathcal{A}_5$. But the $\underline{5}$–cycle $(1\,2\,3\,4\,5)$ is even and not in the kernel (we just calculated where it goes), so the kernel cannot be all of $\mathcal{A}_5$ and hence cannot be anything at all. In other words the action must be faithful.

For more practice, we can of course we can find out where the other adjacent transpositions go. Since

$$(2\,3) = (1\,2\,3\,4\,5)(1\,2)(1\,2\,3\,4\,5)^{-1}$$

we've written $(2\,3)$ in terms of elements that we already know what effect of φ has, so we apply φ to both sides to figure out where $(2\,3)$ goes:

$$\begin{aligned}
\Rightarrow \varphi(2\,3) &= \varphi\Big((1\,2\,3\,4\,5)(1\,2)(1\,2\,3\,4\,5)^{-1}\Big) \\
&= \varphi(1\,2\,3\,4\,5)\varphi(1\,2)\varphi(1\,2\,3\,4\,5)^{-1} && (\varphi \text{ is a gp hom}) \\
&= (2\,6\,5\,4\,3)\Big[(1\,2)(3\,4)(5\,6)\Big](2\,6\,5\,4\,3)^{-1} && (\text{Writing in where gens go}) \\
&= (1\,6)(2\,3)(4\,5). && (\text{Conjugation is replacement})
\end{aligned}$$

Alternatively, like with any homomorphism, we could have gone back to the original definition to figure out where $(2\,3)$ goes. Indeed, for a quick check that we didn't mess things up, note that

$$(2\,3)(1\,2\,3\,4\,5)(2\,3)^{-1} = (1\,3\,2\,4\,5) \in H_6 \Rightarrow H_1 \mapsto H_6$$

which indeed matches the fact that $\varphi(2\,3)$, as appears above, takes the letter 1 to the letter 6.[III]

Similarly,

$$\begin{aligned}
(3\,4) = (1\,2\,3\,4\,5)\,(2\,3)\,(1\,2\,3\,4\,5)^{-1} &\mapsto (2\,3\,4\,5\,6)^{-1}\Big[(1\,6)(2\,3)(4\,5)\Big](2\,3\,4\,5\,6) = (1\,5)(2\,6)(3\,4) \\
(4\,5) = (1\,2\,3\,4\,5)\,(3\,4)\,(1\,2\,3\,4\,5)^{-1} &\mapsto (2\,3\,4\,5\,6)^{-1}\Big[(1\,5)(2\,6)(3\,4)\Big](2\,3\,4\,5\,6) = (1\,4)(5\,6)(2\,3).
\end{aligned}$$

In summary, the monomorphism on this generating set is given by

$$\begin{aligned}
&\varphi\colon \mathcal{S}_5 \to \mathcal{S}_6 \\
&\quad : (1\,2) \mapsto (1\,2)(3\,4)(5\,6) \\
&\quad : (2\,3) \mapsto (1\,6)(2\,3)(4\,5) \\
&\quad : (3\,4) \mapsto (1\,5)(2\,6)(3\,4) \\
&\quad : (4\,5) \mapsto (1\,4)(5\,6)(2\,3).
\end{aligned} \tag{16}$$

EXERCISE 10.43. Another typical generating set of $\mathcal{S}_5$ is

$$\{(1\,2),(1\,3),(1\,4),(1\,5)\},$$

c.f. see Exercise 9.3 back on page 442. Figure out the image of this generating set under φ by

SOLUTION

[III] With this problem it's a matter of taste as to which way you proceed. I recommend you do both so you can really engrain in your head what on Earth a homomorphism does.

(1) Write each as a product of adjacent transpositions

$$\{(1\,2),(2\,3),(3\,4),(4\,5)\}$$

and then map this product to $\mathcal{S}_6$ using (16), and lastly multiply in $\mathcal{S}_6$.
Hint! The basic tricks of multiplying permutations is your best friend here!

(2) Go back to the original definition and see the effect of permuting the H_i's.

(You better get the same answer either way.)

EXERCISE 10.44. Under the map φ from the previous example, prove that

SOLUTION

$$\begin{aligned}
\underline{2}\text{–cycles} &\mapsto \underline{2}\,\underline{2}\,\underline{2}\text{–cycles}\\
\underline{3}\text{–cycles} &\mapsto \underline{3}\,\underline{3}\text{–cycles}\\
\underline{4}\text{–cycles} &\mapsto \underline{4}\text{–cycles}\\
\underline{5}\text{–cycles} &\mapsto \underline{5}\text{–cycles}\\
\underline{2}\,\underline{2}\text{–cycles} &\mapsto \underline{2}\,\underline{2}\text{–cycles}\\
\underline{3}\,\underline{2}\text{–cycles} &\mapsto \underline{6}\text{–cycles.}
\end{aligned}$$

Conclude that the action is faithful.
Hint! Since, like under homomorphism φ, if σ is conjugate to τ (which in $\mathcal{S}_n$ simply amounts to saying that they have the same cycle type) then $\varphi(\sigma)$ is conjugate to $\varphi(\tau)$. Therefore under this particular homomorphism, for instance, you do not need to check that every $\underline{3}\,\underline{2}$–cycle maps to a $\underline{6}$–cycle, you only need to check that this is true for a particular $\underline{2}\,\underline{2}\,\underline{2}$–cycle.

EXERCISE 10.45. Prove that the image of φ from (16) does not meet $\mathcal{S}_3$ (viewed as a subgroup of $\mathcal{S}_6$ in the standard way). Conclude that the 6 elements of $\mathcal{S}_3$ make up the 6 distinct cosets $\mathcal{S}_6/\varphi(\mathcal{S}_5)$.

SOLUTION

EXERCISE$^{\textbf{Hard!}}$ **10.46.** Following the methods of 10.75 prove that $S\left(\text{}^3\right) \cong \mathcal{A}_5$ where, as usual, $S\left(\text{}^3\right)$ is the group of (orientation preserving) symmetries of a Dodecahedron.

SOLUTION

Hint! This problem takes a while to figure out. It's tempting to try and find 5 things which are permuted by $S(\text{}^3)$ but this is incredibly hard to see (if one wikis the right thing, you can see images of 5 inscribed tetrahedron inside of the Dodecahedron which are permuted faithfully by $S(\text{}^3)$ ... but ... right ...). This is part of the reason why I don't want you to attack this problem this way.

For me at least, it is easier to start by writing $S\left(\text{}^3\right)$ as a subgroup of $\mathcal{S}_6$ by permuting the 6 opposite face pairs of the Dodecahedron (one never actually needs to show that this action is faithful or that the subset you choose actually generates $S\left(\text{}^3\right)$ since we will show the image has order 60 by showing that it is isomorphic to $\mathcal{A}_5$!). Then

show that with the suitable choice of a dictionary, this subgroup of $\mathcal{S}_6$ is the same as the image of $\mathcal{A}_5$ inside of $\mathcal{S}_6$ from Example 10.78.

MAPLE EXAMPLE 10.79. We explain how we can use MAPLE allows us to show that $\mathbf{SL}_3(\mathbb{Z}_2)$ is generated by the two matrices

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

(despite the fact that it can't do it directly). Let's use the dictionary on $\mathbb{Z}_2^3 \setminus \{(0,0,0)\}$ given as (reverse) non-zero binary numbers:

$$\begin{array}{lll} 1 = (1,0,0) & 2 = (0,1,0) & 3 = (1,1,0) \\ 4 = (0,0,1) & 5 = (1,0,1) & 6 = (0,1,1) \\ & 7 = (1,1,1). & \end{array}$$

Then with this dictionary, we can see directly that **A** becomes the permutation (2 3)(6 7) and **B** becomes the permutation (1 2 4)(3 6 5). Now plugging in the commands

```
> G:=permgroup(7,{[[1,2,4],[3,6,5]],[[2,3],[6,7]]});
> grouporder(G);
```

returns 168, so the order of the group generated by these two matrices is 168. What's the order of the entire group $\mathbf{SL}_3(\mathbb{Z}_2)$ in question? Either looking ahead to Table 10.1 ahead on page 556, or taking a set of matrices which obviously generate $\mathbf{SL}_3(\mathbb{Z}_2)$, and then using MAPLE, we can see that this is the order of this group. Therefore we have the result. Now that we have this group programmed into our computer we can ask it other things such as what are the possible orders of the elements in this group. Using the command

```
>elo:=proc(g)
return(grouporder(permgroup(7,{g})));
end;
>map(elo,elements(G));
```

(the mini procedure here is due to the fact that you cannot even directly ask MAPLE what the order of an element of a group is!) returns $\{1,2,3,4,7\}$ which means that this group only has elements of these orders. So there is no such matrix which, say, has order 6 or 8.

LEMMA 10.80. *Suppose that we have elements x and y which generate a group G and satisfy*

- $|x| = n$;
- $|y| = 2$;
- $yxy^{-1} = x^{-1}$.

Then if we let $H := \langle y \rangle$, then every $gH \in G/H$ is equal to $x^i H$ for some unique integer $0 \leq i < n$.

PROOF. Notice that $K := \langle x \rangle$ is a set of representatives for the proposed cosets. For starters, if $K \cap H$ were not trivial, then since y has order 2, we would have to have that $y = x^i$ for some integer i. But this would imply that x and y commute; contradicting hypothesis **c**. Armed with this we can prove the two (implied) parts of the lemma.

- We show that there are no repeats in the proposed cosets. That is we show that if $x^iH = x^jH$ for $0 \leq i, j < n$, then $i = j$. Indeed, if $x^iH = x^jH$ were to hold then we would have $x^jx^{-i} = x^{j-i} \in H$. But this is false for $j \neq i$, since otherwise x^{j-i} would be in $H \cap K$.
- We need to show that every coset has that form. Now, the hypothesis guarantee that $K \trianglelefteq G$. Hence we have that $KH = \langle H, K \rangle = G$. As such, for any $g \in G$, $g = x^iy^j$ for integers i and j. But $gH = x^iy^jH = x^iH$ as desired.

■

THEOREM 10.81. *Suppose that we have elements x and y in a group which satisfy*

- $|x| = n$;
- $|y| = 2$;
- $yxy^{-1} = x^{-1}$.

Then $\langle x, y \rangle \cong \mathcal{D}_n$ by an isomorphism determined by

$$x \mapsto \rho_n \qquad y \mapsto \mathbf{r}.$$

PROOF. For ease of notation we do the proof only in the case that $n = 6$ and leave the more general proof to the reader. In this case, we have that

$$G/H = \{H, xH, x^2H, x^3H, x^4H, x^5H\}$$

with no repeats. We use the dictionary

$$\{6 = H, 1 = xH, 2 = x^2H, 3 = x^3H, 4 = x^4H, 5 = x^5H\}.$$

(I could just as easily label H as 0, but this isn't what we've been doing!) Let's see how x and y act on G/H. For starters, we have that

$$x^6 \bullet xH = x^5 \bullet (x^2H) = x^4 \bullet (x^3 \bullet H) = x^3 \bullet (x^4H) = x^2 \bullet (x^5H) = x \bullet H = xH.$$

As such we have that $\rho^\bullet(x) = (1\,2\,3\,4\,5\,6)$. Now the action of y is a little trickier [IV] Notice that in general we have that

$$yx^i = yx^iy^{-1}y = \left(yxy^{-1}\right)^i y = \left(x^{-1}\right)^i y = x^{-1}y.$$

As such we have

$$y \bullet x^iH = yx^iH = x^{-i}yH = x^{-i}H = x^{6-i}H.$$

Therefore, we must have that

$$\rho^\bullet(y) = (1\,5)(2\,4)(3)(6) = (1\,5)(2\,4).$$

Lastly, we show that the subgroup $\langle y \rangle$ is malnormal in G. This is due to the fact that

$$x\langle y \rangle x^{-1} = x\{e, y\}x^{-1} = \{e, xyx^{-1}\} = \{e, x^2y\}$$

Which only meets $\langle y \rangle$ at the identity.

Therefore, any group satisfying the hypothesis of the lemma (with $n = 6$) must be isomorphic to the group of permutations

$$\langle \rho^\bullet(x), \rho^\bullet(y) \rangle = \langle (1\,2\,3\,4\,5\,6), (1\,5)(2\,4) \rangle.$$

But the group $\mathcal{D}_6$ satisfies the hypothesis of the lemma with $\rho = x$ and $\mathbf{r} = y$. Therefore $\mathcal{D}_6$ is isomorphic to this symmetric group via (say ρ') and therefore G and $\mathcal{D}_6$ must be isomorphic to each other via $\rho'^{-1}\rho$, which takes $x \mapsto \rho$ and $y \mapsto \mathbf{r}$. ■

[IV] Of course if y generated a normal subgroup the action would be trivial!

Exercise 10.47. Suppose that

- $|x| = 2n$;
- $yxy^{-1} = x^{-1}$.
- $x^2 = y^2$.

Prove that $\langle x, y \rangle \cong B\mathcal{D}_n$.

SOLUTION

Remark 10.82. Recall that we extended the class of the dihedral groups slightly by defining $\mathcal{D}_2$ as the symmetries of a rectangle which is not a square, $\mathcal{D}_1$ as the symmetries of a line segment. We also define $\mathcal{D}_1$ as the symmetries of a point, and $\mathcal{D}_\infty$ as the symmetries of the integers inside of the real line. The group $\mathcal{D}_\infty$ is called the *infinite dihedral group*.

Of course with these definitions we have that $\mathcal{D}_2 \cong \mathbb{Z}_2 \times \mathbb{Z}_2$, $\mathcal{D}_1 \cong \mathbb{Z}_2$ and $\mathcal{D}_0$ is the trivial group containing only the identity. These definitions are just made to make the statements of theorems a bit nicer and aren't defined for any other reason besides convenience.

Corollary 10.83. *If s and t elements of a group which have order 2, then $\langle s, t \rangle \cong \mathcal{D}_n$ where n is the order of st.* [V]

Proof. Let $y = s$ and $x = st$. Then we have that $\langle x, y \rangle \subseteq \langle s, t \rangle$. OTOH, since $y^{-1}x = s^{-1}x = t$ we have that $\langle s, t \rangle \subseteq \langle x, y \rangle$ and equality holds. Moreover, we have that $|y| = 2$ and $|x| = |st| = n$ by assumption. Lastly, we have that

$$yxy^{-1} = s(st)s^{-1} = s(st)s = (ss)ts = ts = t^{-1}s^{-1} = (st)^{-1} = x^{-1}$$

we can apply the previous theorem to conclude the desired result. ∎

Corollary 10.84. *Any group generated by two elements of order 2 is a dihedral group.*

Proof. This is the previous Corollary since the product of the two generators must have some order. ∎

Rant 10.85. Groups G generated by three (say distinct so we don't have to make special cases) elements of order 2, s_1, s_2, and s_3, are quite interesting. The *triangle groups* fall into this category.[VI] If we let a_1 be the order of s_2s_3, a_2 be the order of s_1s_3, and a_3 be the order of s_1s_2, then we can tesselate $\mathbb{S}^2$ by triangles with angles π/a_i exactly when

$$\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} > 1,$$

we can tesselate $\mathbb{R}^2$ by triangles with angles π/a_i exactly when

$$\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} = 1,$$

[V] This also works when $s = t$ since in this case the product has order 1 (it's the identity!) and the group is isomorphic to $\mathbb{Z}_2 = \mathcal{D}_1$

and we can tesselate the hyperbolic plane by triangles with angles π/a_i exactly when

$$\frac{1}{a_1}+\frac{1}{a_2}+\frac{1}{a_3}<1.$$

In each of these cases G is (isomorphic to) a subgroup of symmetries of this tessellation which acts freely, transitively, with each generator, s_i acting as a reflection across a side opposite the angle π/a_i. In particular G is finite only in the first case. For instance $\mathcal{S}_4$ is generated by $(1\,2)$, $(2\,3)$, and $(3\,4)$. There we have that $a_1=3$, $a_2=2$, and $a_3=3$ ($1/3+1/2+1/3=7/6>1$). Somebody with a basic knowledge of spherical geometry would know that such a triangle would have area $7\pi/6-\pi=\pi/6$.[VII] Since the (surface) area of a sphere is 4π there would be $4\pi/(\pi/6)=24$ triangles in this tessellation. Since the action of $\mathcal{S}_4$ on this tessellation is free and transitive, this corresponds to the fact that there are 24 elements in $\mathcal{S}_4$.

If we know the order of the group in question already, there is no reason to compute the order of their product. This is illustrated in the next example.

Example 10.86. The group $\mathrm{UT}_3(\mathbb{Z}_2)$ is generated by two elements of order 2,

$$\begin{pmatrix}1&1&0\\0&1&0\\0&0&1\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix}$$

and so it is a dihedral group of some order. Since $\mathrm{UT}_3(\mathbb{Z}_2)$ has order 8, the product of these two elements must have order 4 and

$$\mathrm{UT}_3(\mathbb{Z}_2)\cong\mathcal{D}_4.$$

Here's a more general consequence of this. We won't use this result in this book, but it's worth noting since it follows directly from what we have already done.

Corollary 10.87. *Suppose that G is a dihedral group and that φ is a homomorphism. Then $\varphi(G)$ is also a dihedral group.*

Proof. Since $\mathcal{D}_n$ is generated by two elements of order 2, s_1 and s_2, we have that $\varphi(\mathcal{D}_n)=\langle\,\varphi(s_1),\varphi(s_2)\,\rangle$. And each of $\varphi(s_1)$ and $\varphi(s_2)$ have order either 2 or are the identity. If neither are the identity, and $\varphi(s_1)\varphi(s_2)$ has order n, then the image is isomorphic $\mathcal{D}_n$. If exactly one is the identity, then their image is isomorphic to $\mathbb{Z}_2=\mathcal{D}_1$, if both are the identity, then the image is trivial, which is $\mathcal{D}_0$. ∎

10.8.1 Actions on Disjoint Unions

Recall for two sets X and Y the notation

$$X\bigsqcup Y$$

is the disjoint union of X and Y which coincides with the usual union of X and Y when X and Y happen to be disjoint. If X and Y are not disjoint, we think of this disjoint union as the union of "copies of X and Y" which happen to be disjoint. I explain this more rigorously on page 61 way back in Chapter 2.

The "cheapest" way of making an action out of old actions $\star\colon G\circlearrowright X$ and $\dot\star\colon G\circlearrowright Y$ is by considering the action

$$\bullet\colon G\circlearrowright X\bigsqcup Y$$

[VI] Many people call the subgroup of rotational symmetries a triangle group.

[VII] The area of a triangle with angles θ_1,θ_2, and θ_3 is much easier to calculate on the sphere or in hyperbolic space. In the first case the area is equal to how much bigger the sum of the angles are than π, $\theta_1+\theta_2+\theta_3-\pi$ and in the latter case the area is calculated by how much smaller the sum of the angles are than π, $\pi-(\theta_1+\theta_2+\theta_3)$. Of course there cannot even by a formula in the Euclidean case since triangles with the same angles may have different area!

by defining

$$g \bullet z := \begin{cases} g \star z, & \text{if } z \in X \\ g \dot{\star} z, & \text{if } z \in Y \end{cases}.$$

We denote this by writing $\bullet = \star \coprod \dot{\star}$. For example, if $G \bullet x_1, \ldots, G \bullet x_n$ is a decomposition of a G-set X into orbits, then this action defined on

$$G \bullet x_1 \coprod \cdots \coprod G \bullet x_n$$

is the same as the action of G on X. In this way we have that every action can be decomposed into a disjoint union of transitive actions. In other words,

$$\bullet = \bullet_{G \bullet x_1} \amalg \cdots \amalg \bullet_{G \bullet x_n}.$$

Exercise 10.48. Show that if X and Y are non-empty G-sets, then the action of G on $X \amalg Y$ is not transitive. Moreover show that every orbit of this action is contained in either X or Y by showing that for every $x \in X$ and $y \in Y$ that

$$G \bullet x = G \star x \qquad G \bullet y = G \dot{\star} y$$

SOLUTION

Exercise 10.49. Show that

$$\mathrm{Ker}(\bullet) = \mathrm{Ker}(\star) \cap \mathrm{Ker}(\dot{\star}).$$

SOLUTION

These facts proven in the last two exercises helps us send certain groups into $\mathcal{S}_n$ even when there are no malnormal subgroups. In general, since the action $G \circlearrowright G/H$ has kernel equal to the core of H, H°, we have that the action of

$$G \circlearrowright G/H \coprod G/K$$

has kernel equal to the intersection of the core of H and K. In symbols

$$\mathrm{Ker}\left(G \circlearrowright G/H \coprod G/K\right) = H^{\circ G} \cap K^{\circ G}.$$

Example 10.88. We will see that we turn

$$B\mathcal{D}_3 = \left\langle \overset{\chi}{\overset{\shortparallel}{e^{\frac{\pi i}{3}}}}, \mathbf{j} \right\rangle$$

into $\mathcal{S}_n$ as a transitive subgroup iff $n = 12$ and the action is the Cayley one. This is because there are no non-trivial malnormal subgroups of this group. However, we can view this group as a (non-transitive) subgroup of $\mathcal{S}_7$, a vast improvement. We do this by considering the two subgroups H and K we are (resp.)

$$\left\langle \chi^2 \right\rangle \quad \text{and} \quad \langle \mathbf{j} \rangle.$$

Then we have that H has order 3 and K has order 4. Therefore their intersection and their core's intersection must all be trivial, by Lagrange's theorem. So

$$B\mathcal{D}_3 \circlearrowright X$$

is faithful, where we are letting X denote the disjoint union of the two quotients:

$$B\mathcal{D}_3/_H \bigsqcup B\mathcal{D}_3/_K.$$

Let's compute this action. For starters,

$$B\mathcal{D}_3/_H = \{H, \mathbf{j}H, -H, -\mathbf{j}\,H\}$$

and

$$B\mathcal{D}_3/_K = \{K, \chi K, \chi^2 K\}.$$

Therefore

$$X = \{H, \mathbf{j}\,H, -H, -\mathbf{j}\,H, K, \chi K, \chi^2 K\}.$$

Next, let's start to move things around:

$$H = \mathbf{j}^4 \bullet H = \mathbf{j}^3 \bullet jH = \mathbf{j}^2 \bullet \mathbf{j}^2 H = \mathbf{j} \bullet \mathbf{j}^3 H = H$$
$$K = \mathbf{j} \bullet K$$

and

$$\begin{aligned}
\mathbf{j} \bullet \chi K &= \mathbf{j} \bullet \chi \mathbf{j}^{-1} K \quad (\text{Since } \mathbf{j}^{-1} \in K = \langle \mathbf{j} \rangle \Rightarrow \mathbf{j}^{-1} K = K) \\
&= \mathbf{j}\chi\mathbf{j}^{-1} K \\
&= \chi^{-1} K \\
&= \chi^5 K \qquad (\text{Since } \chi = \chi_6 \text{ has order 6}) \\
&= \chi^2 \chi^3 K \\
&= \chi^2 K \qquad (\text{Since } \chi^3 = -1 = \mathbf{j}^2 \in K = \langle \mathbf{j} \rangle).
\end{aligned}$$

Therefore, $\mathbf{j} \bullet \chi^2 K = \chi K$ (this is the only thing that doesn't already have somebody mapping to it!) At the same time

$$\begin{aligned}
\chi \bullet H &= \chi \bullet \chi^2 H \quad (\text{Since } \chi^2 \in H = \left\langle \chi^2 \right\rangle) \\
&= \chi^3 H \\
&= -1H.
\end{aligned}$$

And since $\chi^2 \in H$ we have that $\chi^2 H = \chi \bullet (-H) = H$. At the same time, playing with the cosets of H instead we see that

$$\begin{aligned}
\chi \bullet \mathbf{j}H &= \chi \mathbf{j} H \\
&= \mathbf{j}\mathbf{j}^{-1}\chi\mathbf{j}H \\
&= \mathbf{j}\chi^5 H \\
&= \mathbf{j}\chi^2(-H) \\
&= \mathbf{j}(-H) \qquad (\text{Since } \chi^2 \text{ is in the normal subgroup } H) \\
&= -\mathbf{j}H.
\end{aligned}$$

And again, since $\chi^2 \in H$ and H is normal, we must have that $\chi^2 \bullet \mathbf{j}\, H = \mathbf{j}\, H$. Lastly, the action of $\mathbf{j}$ on K is easy to find since $-1 \in K$ implies that

$$\chi^3 \bullet K = K = \chi^2 \bullet \chi K = \chi \bullet \chi^2 K = \chi^3 K.$$

So if we define the dictionary

$$\{\overset{1}{H}, \overset{2}{\mathbf{j}\, H}, \overset{3}{-H}, \overset{4}{-\mathbf{j}H}, \overset{5}{K}, \overset{6}{\chi K}, \overset{7}{\chi^2 K}\}$$

we have that $\mathbf{j}$ acts as $(1\,2\,3\,4)(6\,7)$ and χ acts as $(1\,2)(3\,4)(5\,6\,7)$. So we have an injective (since the action is faithful) homomorphism,

$$B\mathcal{D}_3 \to \mathcal{S}_7$$

whose image is $\langle (1\,2\,3\,4)(6\,7), (1\,2)(3\,4)(5\,6\,7) \rangle$. In conclusion,

$$B\mathcal{D}_3 \cong \langle (1\,2\,3\,4)(6\,7), (1\,2)(3\,4)(5\,6\,7) \rangle.$$

Remark 10.89. Its worth noting that the letters $\{1, 2, 3, 4\}$ and $\{5, 6, 7\}$ make up different orbits in the action from this last example! We knew this would happen because these numbers correspond to cosets of different groups

10.8.2 Classification of Groups of Order 8

Later in Corollary 13.54 coming up on page 709 we will see there are 3 abelian groups of order, namely the 3 obvious ones:

$$\mathbb{Z}_8 \quad \mathbb{Z}_4 \times \mathbb{Z}_2 \quad \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2.$$

Therefore for now we focus on the non-abelian case.

Lemma 10.90. *Let G be a group which satisfies $g^2 = e_G$ for every $g \in G$. Then G is abelian.*

Proof. Let a and b be arbitrary elements of G.

$$\begin{aligned} aba^{-1}b^{-1} &= abab && \text{(Since } a \text{ and } b \text{ have order 2)} \\ &= (a\,b)^2 \\ &= e_G. && \text{(Since } ab \in G \Rightarrow (ab)^2 = e_G) \end{aligned}$$

Therefore

$$aba^{-1}b^{-1} = e_G \Rightarrow aba^{-1} = b \Rightarrow ab = ba.$$

Since a and b are arbitrary, G is abelian. ∎

Armed with this, we let G be an arbitrary non-abelian group with 8 elements (i.e. of order 8). Let x be an element of largest order. By Lagrange's theorem, we know that this order must be 2, 4, or 8 (can't be 1 since this would mean that G had order 1). If this order is 2, then by the lemma G would be abelian which we are not considering. If on the other hand this order is 8, then we would have that $G = \langle x \rangle$ which again is abelian. Therefore, we are left with the case that $|x| = 4$. Now let b be an element of G which is not in $\langle x \rangle$. Such an element exists since $\langle x \rangle$ has order 4 but our G has order 8. Since $e_G \in \langle x \rangle$ we must have that $b \neq e_G$ and, as such, has order 2 or 4 (again, 8 is ruled out).

Next, we note that $\langle x, y \rangle = G$. This is because the order of the first group, $\langle x, y \rangle$, is a divisor of 8 but is strictly larger than 4 (since it contains $\langle x \rangle$ but does not equal this since it contains y which is not in there). The only such divisor is 8, so $\langle x, y \rangle$ has order 8. Since it is contained inside a group of order 8, G, it must be all of G.

Now we observe that $\langle x \rangle$ has index $8/4 = 2$ in G, and it is therefore normal in G by Theorem 10.61. In particular, we must have that

$$yxy^{-1} \in \langle x \rangle \Rightarrow yxy^{-1} = x^i \text{ for some } i \in \mathbb{Z}.$$

If $i = 1$, we would have that $yxy^{-1} = x$ of $yx = xy$ and yet again we would have that G is abelian (since it would imply that G is generated by two commuting elements). Of course if $i = 0$ we would have that $x = e$, again, a load of crap since $|x| = 4$. The next case to rule out is the case that $i = 2$. But if i did indeed equal 2 we would have that

$$y^2xy^{-2} = y(yxy^{-1})y^{-1} = yx^2y^{-1} = \left(yxy^{-1}\right)^2 = \left(x^2\right)^2 = x^4 = e_G$$

which again leads us to our crap statement that $x = e_G$ (by "unconjugating" by y^2). Therefore, we must have that $i = 3 \equiv_4 -1$. We are therefore left with two cases

(1) The order of y is 2. Then we must have that $G \cong \mathcal{D}_4$.

(2) The order of y is 4. In this case we want to show that G is the quaternion group. Therefore we only need to show that $y^2 = x^2$. Now, since $\langle x \rangle$ has index 2 in G we must have that $y^2 \langle x \rangle = y^2 \bullet \langle x \rangle = \langle x \rangle$ so that $y^2 = x^j$ for some integer j between 0 and 3. If $j = 0$, then we would have that $y^2 = e_G$ contradicting the assumption that $|y| = 4$. If $j = 1$ or $j = 3$, then we would have that y^2 would have order 4. But this would imply that y would have order 8, another contradiction. The only case left is $j = 2$, which would imply that $x^2 = y^2$, which is all we needed.

Adding in what we will prove about abelian groups in Chapter 13, we have

Every abelian group of order 8 is isomorphic to exactly one of the following three groups:

(1) $\mathbb{Z}_8$

(2) $\mathbb{Z}_4 \times \mathbb{Z}_2$

(3) $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$

Every non–abelian group of order 8 is isomorphic to exactly one of the following two groups:

(1) $\mathcal{D}_4$

(2) Q_8

10.9 The Orbit Stabilizer Theorem

One of the most important ways that applications of group actions come up is via Orbit-Stabilizer theorem (**OST**). This theorem gives us a simple equation which relates the size of an orbit to the size of the stabilizer of any point in the orbit. Many results from combinatorics are really applications of this kick butt theorem.

10.9.1 Statement and Proof

In every branch of mathematics (formally called a category in graduate mathematics), one always has to ask what it means for two things to be equivalent. An equivalence is typically stated as a bijection which preserves the relevant properties. For instance, an equivalence in groups is an isomorphism, i.e. a bijection which preserves the group's operation. We therefore pose the question of what it means for two G-sets to be equivalent (the wording is a G-equivariant equivalence). As you would guess G-equivariant

equivalence preserves action diagrams. However, just like saying that to groups are isomorphic has its flaws, the definition in terms of action diagrams is not as good as it should be. For one thing, the action diagram depends on the generating set involved. Therefore, it depends on a choice. Just like a group's multiplication table depends on a choice of the way in which you write the elements of the group down.

Definition 10.91. A *G-equivariant equivalence* of two G-sets $\bullet\colon G \circlearrowright X$ and $\star\colon G \circlearrowright Y$ is a bijection

$$\varphi\colon X \to Y$$

which satisfies $\varphi(g \bullet x) = g \star \varphi(x)$.

Exercise 10.50. Define a relation $\approx$ on G-sets by $X \approx Y$ iff there is a G-equivariant equivalence $\varphi\colon X \to Y$. Show that $\approx$ is an equivalence relation.
Hint! Follow the usual outline; show that the identity is a G-equivariant equivalence, then show the inverse of a G-equivariant equivalence is a G-equivariant equivalence, and finally show that the composition of two G-equivariant equivalences is a G-equivariant equivalence.

SOLUTION

Exercise 10.51. Show that if

$$\varphi\colon X \to Y$$

is a G-equivariant equivalence, then if we replace the vertices of the action diagram of $G \circlearrowright X$ by the vertices of $G \circlearrowright Y$ the two diagrams line up. State a reasonable converse and prove it.

SOLUTION

Theorem 10.92 (OST)**.** *Suppose that a group G acts transitively on a set X and let $x \in X$. Then we have a G-equivariant equivalence*

$$G/G_x \to X$$

defined by $gG_x \mapsto g \bullet x$.

Proof. We have a function defined on cosets, so we need to make sure it's well defined before proving that it's G-equivariant and bijective. Let's call the supposed function φ (for now φ is just a relation).

(1) We need to show that written function is well defined. Suppose that $a \equiv_{G_x} b$ (that is $aG_x = bG_x$). Then we need to show that $\varphi(a) = \varphi(b)$. But, since $a \equiv_{G_x} b$ we have that $b^{-1}a \in G_x$. However, this means that $b^{-1}a \bullet x = x$. But, this means (bringing the b to the other side) that $a \bullet x = b \bullet x$. But the left hand side of this equation is $\varphi(a)$ and the right hand side is $\varphi(b)$.

(2) Next we show that φ preserves the action. To that end, let $aH \in G/H$ be arbitrary and let $g \in G$. Then we have that

$$\varphi(g \bullet (aG_x)) = \varphi(gaG_x) = ga \bullet x = g \bullet (a \bullet x) = g \bullet (\varphi(aG_x)).$$

(3) Now we show that φ is onto. But since the action of G on X is transitive, for any $\overline{x} \in X$ there exists a $g \in G$ so that $g \bullet x = \overline{x}$. But this means that $\varphi(gG_x) = g \bullet x = \overline{x}$ and so φ is surjective.

(4) Lastly, we need to show that φ is one-to-one. To this end, suppose that $\varphi(aG_x) = \varphi(bG_x)$. We need to show that $aG_x = bG_x$. However, since

$$\varphi(aG_x) = a \bullet x = b \bullet x = \varphi(bG_x)$$

we must have that $b^{-1}a \bullet x = x$. That is $b^{-1}a \in G_x$ which finally implies that $aG_x = bG_x$ as we wanted. ∎

COROLLARY 10.93 (OST II). *Let $x \in X$ and suppose that G acts on X. Then, the following*

$$|G| = \#(G \bullet x) \cdot |G_x|.$$ ∎

If $G \circlearrowright X$ is transitive, then for any point in x we have $G \bullet x = X$. Therefore we have

COROLLARY[2] 10.94 (OST III). *Suppose $G \circlearrowright X$ is transitive. Then, for any $x \in X$*

$$|G| = \#X \cdot |G_x|.$$ ∎

REMARK 10.95. Since we have that $G_{g \bullet x} \cong G_x$ via Theorem 10.34, it shouldn't surprise you too much that these formula do not depend on which point we pick for x.

Either of these Corollaries are also often called the orbit stabilizer theorem. At its very least it says that the size of the orbit of any point under any action divides the order of the group. For instance, if we look at the action of G on itself by conjugation, an orbit is what we called a conjugacy class before. This says that the size of any conjugacy class divides the order of the group. This is a special case (a particular action) of this special case (just saying that the order divides it) of this special case (the corollary) of the theorem and it is not even intuitive.

10.9.2 First Applications

There are literally so many examples of the use of the **OST** that it's difficult to know where to start! This theorem holds in such a general context (all group actions) and its statement is so precise that the consequences are nearly unlimited!

EXAMPLE 10.96. A simple application of **OST** allows us to calculate something that we already know, but probably have never seen derived before (and certainly not like this!); namely that the number of ways to choose m things from n things is

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}.$$

In terms of sets, we are looking at the set of m element subsets of an n element set. Now $\mathcal{S}_n$ acts on these m element sets and this action is clearly transitive. If we take $\overline{n}$ as our n-element set, the stabilizer of $\overline{m} \subseteq \overline{n}$ is

$$\mathcal{S}_{\{1,\dots,m\}} \times \mathcal{S}_{\{m+1,\dots,n\}} \cong \mathcal{S}_m \times \mathcal{S}_{n-m}.$$

Therefore by **OST** we have proven the well known result

$$\#\mathcal{P}_m(\overline{n}) = \frac{|\mathcal{S}_n|}{|\mathcal{S}_m \times \mathcal{S}_{n-m}|} = \frac{n!}{m!\,(n-m)!}.$$

EXERCISE 10.52. Without looking or googling, use the fact that the symmetry group of a soccer ball has the same number (60) of symmetries as an icosahedron to compute

(1) How many pentagon faces are there in a soccer ball?

(2) How many hexagon faces are there in a soccer ball?

(3) How many vertices are there in a soccer ball?

(4) How many square faces are there in a soccer ball?

SOLUTION

Hint! See Exercise C.19 coming up on page 796.

EXAMPLE 10.97. **(1)** We have already shown that $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ acts transitively on the p^2-1 elements of

$$\mathbb{Z}_p^2 \setminus \left\{\begin{bmatrix}[0]_p\\[0]_p\end{bmatrix}\right\}$$

and $\begin{bmatrix}[1]_p\\[0]_p\end{bmatrix}$ has stabilizer the set of p matrices in $\mathrm{UT}_2\left(\mathbb{Z}_p\right)$, i.e. the set of matrices of the form

$$\begin{pmatrix}[1]_p & [a]_p\\[0]_p & [1]_p\end{pmatrix}.$$

Hence $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ has $(p^2-1)(p) = (p+1)\,p\,(p-1)$ elements. We record this as an example/theorem:

EXAMPLE /THEOREM 10.98. For any positive prime integer,

$$\left|\mathbf{SL}_2\left(\mathbb{Z}_p\right)\right| = (p+1)\,p\,(p-1).$$

■

Let's make a table to get a better sense of the numbers:

p	2	3	5	7	11	13	17
$\left\|\mathbf{SL}_2\left(\mathbb{Z}_p\right)\right\|$	6	24	120	336	1320	2184	4896

.

(2) Figuring out the order of $\mathbf{SL}_2\left(\mathbb{Z}_n\right)$ when n is not a prime, but is a composite, is more difficult. The reason why is, in this case, the action of the group is no longer transitive. Let's consider the case that $n=4$. We still have that the stabilizer of $([1]_4, [0]_4)^\top$ consists of the 4 matrices of the form

$$\begin{pmatrix}[1]_4 & [a]_4\\[0]_4 & [1]_4\end{pmatrix}.$$

(as $[a]_4$ goes through the four elements of $\mathbb{Z}_4$). Now we need to figure out what the orbit of $([1]_4, [0]_4)^\top$ is. For starters, the orbit cannot contain any of the elements

$$\begin{bmatrix}0\\0\end{bmatrix}\ \begin{bmatrix}0\\2\end{bmatrix}\ \begin{bmatrix}2\\0\end{bmatrix}\ \begin{bmatrix}2\\2\end{bmatrix}$$

(all numbers understood to be number mod 4) since each would lead to a determinant which is a multiple of 2 in $\mathbb{Z}_4$, hence not equal to 1. Now each of the $16-4$ remaining elements of $\mathbb{Z}_4 \times \mathbb{Z}_4$ have an entry which is equal to either 1 or 3. In particular, each must have an entry which is invertible in $\mathbb{Z}_4$. This can easily be used to construct a matrix whose first column is that given column. Thus the orbit of the action of $\mathbf{SL}_2(\mathbb{Z}_4)$ consists of 12 points. Therefore, by the **OST**, $\mathbf{SL}_2(\mathbb{Z}_4)$ contains $2 \cdot 12 = 48$ elements.

(3) Next let's try to find the order of $\mathbf{SL}_3\left(\mathbb{Z}_p\right)$. Here, we have an action of this group on the set of p^3-1 elements of $\mathbb{Z}_p^3 \setminus \{(0,0,0)\}$. The stabilizer is the set of matrices of the form

$$\begin{pmatrix} 1 & x & y \\ 0 & a & b \\ 0 & c & d \end{pmatrix}$$

with $(ad-bc)=1$. Now the 2 by 2 submatrix given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

gives us a copy of $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ sitting inside of this stabilizer while the x and y aren't restricted and can take only any of the p^2 possibilities. Hence the stabilizer has order $p^2 \cdot ((p+1)\,p\,(p-1))$ and so by the **OST** we have that

$$\left|\mathbf{SL}_3\left(\mathbb{Z}_p\right)\right| = (p^3-1) \cdot p^2 \cdot (p+1)\,p\,(p-1) = p^3 \cdot (p^3-1) \cdot (p^2-1) = p^8 - p^6 - p^5 + p^3$$

elements, see Table 10.1.

p	2	3	5	7
$\left\|\mathbf{SL}_3\left(\mathbb{Z}_p\right)\right\|$	168	5616	372,000	5,630,688
$=$	$2^3 \cdot 3 \cdot 7$	$2^4 \cdot 3^3 \cdot 13$	$2^5 \cdot 3 \cdot 5^3 \cdot 31$	$2^5 \cdot 3^3 \cdot 7^3 \cdot 19$

Table 10.1: Order of $\mathbf{SL}_3\left(\mathbb{Z}_p\right)$.

Let $\mathcal{P}$ be a set consisting of an even number, $2k$, of people. Define a hand holding to be a situation where every person in $\mathcal{P}$ holds the hand of exactly one other person in $\mathcal{P}$.

SOLUTION

EXERCISE 10.53.

(1) Prove that there are a total of

$$\frac{(2n)!}{2^n \cdot n!}$$

possible hand holdings.

(2) For the case that $n=3$ explain the relationship between this and the symmetries of a cube.

EXERCISE 10.54. Suppose $\mathbf{A}$ is a 3 by 3 matrix with entries in $\mathbb{Z}_3$. Show that $\mathbf{A}^5 = \mathbf{I}_3$ implies that $\mathbf{A} = \mathbf{I}_3$, the 3 by 3 identity matrix.

SOLUTION

EXAMPLE 10.99.

(1) Let G act on itself my conjugation. Then we have that for any $x \in G$, that $G_x = C_G(x)$ the centralizer of x in G. Moreover the orbit of x is the conjugacy class of x. Therefore, the orbit stabilizer theorem guarantees us that

PROPOSITION 10.100.

$$|G| = \#(\text{cc}_G(x)) \cdot |C_G(x)|.$$

■

(2) Since we have that $\mathrm{N}_G(H)$ is the stabilizer of the action $G \circlearrowright \{\text{Subgroups of G}\}$ the **OST** guarantees us that the number conjugate subgroups of H in G, $\#(\text{cc}_G(H))$ is equal to $|G|/|\mathrm{N}_G(H)|$.

PROPOSITION 10.101.

$$|G| = \#(\text{cc}_G(H)) \cdot |\mathrm{N}_G(H)|.$$

■

This will become relevant when we discuss the Sylow theorems because we will see that for many subgroups of a group of a certain fixed order, all groups of that same order will be conjugate.

(3) Here is another formula that we can derive using **OST**. It also illustrates an application of **OST**. Recall that, by definition, the HK which shows up is simply the set of all products $\{hk : h \in H, k \in K\}$.

PROPOSITION 10.102.

$$\#(H \cdot K) = \frac{|H| \cdot |K|}{|H \cap K|}$$

PROOF. The group $H \times K$ acts on the set HK via $(h,k) \bullet x := hxk^{-1}$.

- This $\bullet$ is a group action since

(a) For any $x \in X$

$$\begin{aligned}(e_H, e_K) \bullet x &= e_H x e_K^{-1} \\ &= e_G x e_G^{-1} \\ &= x.\end{aligned}$$

and the identity of $H \times K$ acts as the identity.

(b) For all $h, h' \in H$, $k, k' \in K$ and $x \in X$

$$\begin{aligned}(h,k) \bullet \Big((h',k') \bullet x\Big) &= (h,k) \bullet (h' x k'^{-1}) \\ &= h(h' x k'^{-1})k^{-1} \\ &= (hh')x(kk')^{-1} \\ &= (hh', kk') \bullet x \\ &= \Big((h,k) \cdot (k,k')\Big) \bullet x\end{aligned}$$

and $\bullet$ is associative. Hence it is an action.

- Since $(h, k^{-1}) \bullet e_G = h \cdot (k^{-1})^{-1} = hk$ this action is transitive.
- Next, the stabilizer of e_G is

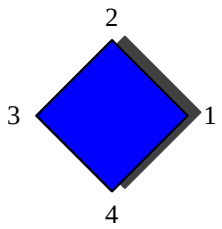
$$\begin{aligned}\{(h,k) : (h,k) \bullet e_G = e_G\} &= \{(h,k) : hek^{-1} = e\} \\ &= \{(h,k) : h = k\} \\ &= \{(h,h) : h \in H \cap K\}\end{aligned}$$

and this last set is clearly in bijection with $H \cap K$ via a $h \mapsto (h,h)$. Therefore the orbit has $\#(HK)$ elements in it and the stabilizer has $|H \cap K|$ elements in it and the **OST** gives the result. ∎

Example 10.103. For an example of some insight into the formula from Proposition **??**, fix a copy of $\mathcal{D}_4 \cong \widetilde{\mathcal{D}_{41}} \leqslant \mathcal{S}_4$. For instance, we could take

$$\widetilde{\mathcal{D}_{41}} = \langle (1\,2\,3\,4), (1\,3) \rangle.$$

which would could get by labeling the square as



By Lagrange's theorem, we know that

$$8 = |H| \;\Big|\; \left|N_{\mathcal{S}_4}(H)\right| \;\Big|\; |\mathcal{S}_4| = 24.$$

Since 8 is a maximal divisor of 24 – since 24/8 is prime – we must have that $N_G(H)$ has order 8, and hence would be H, or it would have order 24, and hence it would be all of $\mathcal{S}_4$. However, we know that the latter happens iff H is normal in $\mathcal{S}_4$. However, this is clearly not the case since

$$(1\,2\,3\,4) \in H, \qquad \text{but} \qquad (1\,2)(1\,2\,3\,4)(1\,2)^{-1} = (2\,1\,3\,4) = (1\,3\,4\,2)$$

is not a symmetry of the square as labeled above (if some jerk permuted the vertices of the labeled square like this they'd rip the poor square to pieces!). Therefore, we must have that $N_G(H)$ has order 8 and hence is just H itself. Hence there are $3 = 24/3$ conjugate copies of $\mathcal{D}_4$ inside of $\mathcal{S}_4$. Indeed, if we think of $\mathcal{S}_4$ as $S\left(\blacklozenge^3\right)$, then these three copies can be seen as the stabilizers of the 3 pairs of opposite faces ... or even better as the intersection of the 3 coordinate hyper planes with the cube or even better the three coordinate axis. (i.e. the x,y-plane, the x, z-plane and the y,z-plane).

The permutation of these three planes (which by **OST** is the same as the conjugation of the 3 copies of $\mathcal{D}_4$) is exactly what gives rise to the homomorphism of $\mathcal{S}_4 \to \mathcal{S}_3$.

Maple Example 10.104. MAPLE does not have a command which calculates the number of conjugate copies of one group inside of a group directly, however it does compute the order of the normalizer subgroup. For instance let G be the group

$$\left\langle (1\,2\,\overset{x}{3}\,4\,5\,6), (1\,8\,5\,9)(2\,\overset{y}{7}\,4\,10)(3\,6) \right\rangle$$

inside of $\mathcal{S}_{10}$ and let's consider the cyclic subgroup $\langle x \rangle$ sitting inside of G which, of course, we call H. We start this by defining G and the subgroup H:

```
> G:=permgroup(10, {[[1,2,3,4,5,6]], [[1,8,5,9],[2,7,4,10],[3,6]]});
> H:=permgroup(10, {[[1,2,3,4,5,6]]});
```

We need to compute the number of elements in G and the number of elements in the normalizer, $N_G(H)$ to figure out the number of conjugacy classes of the subgroup H. The command

```
> grouporder(G);
```

returns 100. And using the command

```
> grouporder(normalizer(G,H));
```

returns the number 50. Therefore there are $2 = 100/50$ conjugate subgroups of $\langle x \rangle = H$. {Why not do this by hand to compute the order of the group G? }

Exercise 10.55. Let p a prime number and n any positive integer. Prove

(1) $\left|\mathbf{SL}_n\left(\mathbb{Z}_p\right)\right| = p^{\binom{n}{2}} \prod_{i=2}^{n} \left(p^i - 1\right)$ I

(2) $\left|\mathbf{GL}_n\left(\mathbb{Z}_p\right)\right| = p^{\binom{n}{2}} \prod_{i=1}^{n} \left(p^i - 1\right)$

Hint!

SOLUTION

- Generalize the proofs given already in the special cases that $n = 2$, $n = 3$, and $n = 4$ by you.
- You only need to prove one of these, and then use the **KIT**, Theorem 10.107 back on page 560, applied to the homomorphism $\det : \mathbf{GL}_n\left(\mathbb{Z}_p\right) \to \mathbb{Z}_p^\times$ (why is this map surjective?), whose kernel is, by definition, $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$.

Remark 10.105. Note that we could have rewritten the quantity in the exercise as

$$\left|\mathbf{GL}_n\left(\mathbb{Z}_p\right)\right| = \prod_{i=0}^{n-1}\left(p^n - p^i\right) = (p^n - p^{n-1})(p^n - p^{n-2})(p^n - p^{n-3})\cdots(p^n - p)(p^n - 1).$$

For instance, for $n = 4$,

$$\left|\mathbf{GL}_4\left(\mathbb{Z}_p\right)\right| = (p^4 - p^3)(p^4 - p^2)(p^4 - p)(p^4 - 1)$$

and for $n = 2$

$$\left|\mathbf{GL}_2\left(\mathbb{Z}_p\right)\right| = (p^2 - p)(p^2 - 1).$$

Notice, by this exercise, that we have that

$$\left|\mathbf{SL}_n\left(\mathbb{Z}_p\right)\right| = \left|\mathbf{SL}_{n-1}\left(\mathbb{Z}_p\right)\right| \cdot p^{n-1} \cdot (p^n - 1).$$

Thus, if a prime $q \neq p$ divides $\left|\mathbf{SL}_n\left(\mathbb{Z}_p\right)\right|$, and it does not divide $\left|\mathbf{SL}_{n-1}\left(\mathbb{Z}_p\right)\right|$, then it must divide $p^n - 1$. Moreover, since we are assuming it does not divide $\left|\mathbf{SL}_{n-1}\left(\mathbb{Z}_p\right)\right|$, it cannot divide any of

$$(p^{n-1} - 1), (p^{n-2} - 1), \cdots, (p - 1), \text{ or } (p + 1)$$

since all of these factors show up in $\left|\mathbf{SL}_{n-1}\left(\mathbb{Z}_p\right)\right|$. Therefore we have proven:

Theorem 10.106. *There exists an element of order q, q a prime, in $\mathbf{SL}_n\left(\mathbb{Z}_p\right)$ and not in $\mathbf{SL}_{n-1}\left(\mathbb{Z}_p\right)$ iff $[\,p\,]_q$ has order n in $\mathcal{U}(\mathbb{Z}_q)$.*

Proof. We are essentially done, I leave it to you to tie everything together. ∎

Exercise 10.56. What is the of $\mathbf{SL}_4(\mathbb{Z}_2)$ and $\mathbf{SL}_4(\mathbb{Z}_3)$?

SOLUTION

Theorem 10.107 (The Kernel/Image Theorem, **KIT**). *Let φ be any homomorphism with domain G. Then*

$$|G| = |\operatorname{Ker}(\varphi)| \cdot |\operatorname{Im}(\varphi)|.$$

[1] Where as usual $\binom{n}{2}$ means n choose 2, i.e. it's the number

$$\frac{n \cdot (n-1)}{2}.$$

Proof. We found an action that, in Equation (3) and Equation (4) backed on 489 which had $\operatorname{Ker}(\varphi)$ the stabilizer of a point whose orbit was $\operatorname{Im}(\varphi)$. The theorem now follows at once from **OST**. ∎

This implies that the order of G divided by the order of the image of any homomorphism is an integer (it's the order of the kernel, i.e. it's equal to the number of elements of some finite set, which is always a ⟨positive⟩ integer.) In particular we have the following corollary.

Corollary 10.108. *For a finite group G, the order of the image of any homomorphism with domain G divides G.* ∎

Remark 10.109. Of course, **KIT** also implies that the order of the kernel of any homomorphism also divides G, but since the kernel of any such homomorphism is a subgroup of G, Lagrange's theorem already takes care of this case.

Corollary 10.110. *Let G and H be two finite groups. If $|G|$ and $|H|$ are relatively prime, then the only homomorphism from G to H is the trivial homomorphism.*

Proof. Let $\varphi\colon G \to H$ be a homomorphism. Let's denote the order of $\operatorname{Im}(\varphi)$ by m. By the previous corollary, we know that m divides the order of G. Moreover, since $\operatorname{Im}(\varphi) \leqslant H$, Lagrange's theorem implies that m divides the order of H. But since $|G|$ and $|H|$ are assumed to be relatively prime, we must have that $m = 1$, i.e. $|\operatorname{Im}(\varphi)| = 1$. But since $\operatorname{Im}(\varphi)$ is a subgroup of H, this means that the one element it (the image of φ) contains must be the identity (of H). In symbols,

$$\{\varphi(g) : g \in G\} = \{e_H\}.$$

That is $\varphi(g) = e_H$ for every $g \in G$, i.e. φ is trivial. ∎

The following exercise shows that the order of the kernel of a homomorphism needn't have anything to do with the order of the codomain. For instance, there is no requirement that the order of the kernel divide the order of the codomain, even for a surjective homomorphism.

Exercise 10.57. Find an example of a homomorphism of groups, $\varphi\colon G \to H$, with $|H|$ and $\operatorname{Ker}(\varphi)$ relatively prime.
Hint! At some point, did we define a surjective homomorphism $\mathcal{A}_4 \to \mathbb{Z}_3$?

SOLUTION

Example 10.111. Let n and m be arbitrary positive integers. If we consider the homomorphism

$$\det\colon \mathbf{GL}_m(\mathbb{Z}_n) \to \mathcal{U}(\mathbb{Z}_n)$$

then, since $\mathcal{U}(\mathbb{Z}_n)$ is a group (under multiplication) with identity $1 = ([\,1\,]_n)$, the kernel is the special linear group (in slightly more detail)

$$\begin{aligned}\operatorname{Ker}(\det) &= \{\mathbf{A} \in \mathbf{GL}_m(\mathbb{Z}_n) : \det \mathbf{A} = e_{\mathcal{U}(\mathbb{Z}_n)}\}\\ &= \{\mathbf{A} \in \mathbf{GL}_m(\mathbb{Z}_n) : \det \mathbf{A} = 1\}\\ &= \mathbf{SL}_m(\mathbb{Z}_n).\end{aligned}$$

Moreover for any $[\,a\,]_n \in \mathbb{Z}_n$, the diagonal matrix which is the identity except for the upper left entry, which is $[\,a\,]_n$,

$$\begin{pmatrix} [\,a\,]_n & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \cdots & \vdots\\ 0 & 0 & \cdots & 1\end{pmatrix},$$

has determinant $[\,a\,]_n$. Therefore the map $\det\colon \mathbf{GL}_m(\mathbb{Z}_n) \to \mathcal{U}(\mathbb{Z}_n)$ is a surjective homomorphism with kernel $\mathbf{SL}_m(\mathbb{Z}_n)$. Therefore, by **KIT**, we have

$$|\,\mathbf{GL}_m(\mathbb{Z}_n)\,| = |\,\mathcal{U}(\mathbb{Z}_n)\,| \cdot |\,\mathbf{SL}_m(\mathbb{Z}_n)\,|\,.^{\text{II}}$$

In particular, for p a prime, we have

$$|\,\mathbf{GL}_m(\mathbb{Z}_n)\,| = (p-1) \cdot |\,\mathbf{SL}_n(\mathbb{Z}_n)\,|\,.$$

Example 10.112. Let's take a look at more abstract matrix groups then $\mathbf{GL}_n\left(\mathbb{Z}_p\right)$ and $\mathbf{SL}_n\left(\mathbb{Z}_p\right)$ that we've been dealing with: the $\mathbf{Sp}_{2n}\left(\mathbb{Z}_p\right)$'s from Exercise 10.28 back on page 515. For concreteness, I only look at the case that $n = 2$ (so 4 by 4 matrices), and I leave the more general case to you in Exercise 10.58. In the stated exercise you prove that the action $\mathbf{Sp}_4\left(\mathbb{Z}_p\right) \circlearrowright \mathbb{Z}_p^4 \setminus \{\mathbf{0}\}$ is transitive. Therefore, since there are $p^4 - 1$ elements in $\mathbb{Z}_p^4 \setminus \{\mathbf{0}\}$ by **OST**, we have for any $\mathbf{v} \in \mathbb{Z}_p^4$

$$\left|\mathbf{Sp}_4\left(\mathbb{Z}_p\right)\right| = (p^4-1) \cdot \left|\mathbf{Sp}_4\left(\mathbb{Z}_p\right)_{\mathbf{v}}\right|. \tag{17}$$

In particular,

$$\left|\mathbf{Sp}_4\left(\mathbb{Z}_p\right)\right| = (p^4-1) \cdot \left|\mathbf{Sp}_4\left(\mathbb{Z}_p\right)_{\mathbf{e}_1}\right|. \tag{18}$$

This gives us hope. However, it is not immediately clearly what the size of the stabilizer of a particular vector (let's take $\mathbf{e}_1$ for ease) is. Our trick will be to double down and use the orbit stabilizer theorem again to take a look at the stabilizer of $\mathbf{e}_2$ in $\mathbf{Sp}_4\left(\mathbb{Z}_p\right)_{\mathbf{e}_1}$. This time we'll have better luck computing the stabilizer.[III] We are looking at symplectic matrices of the form

$$\begin{pmatrix} 1 & 0 & B_{11} & B_{21} \\ 0 & 1 & B_{21} & B_{22} \\ 0 & 0 & A_{11} & A_{21} \\ 0 & 0 & A_{21} & A_{22} \end{pmatrix}$$

or written as a block matrix:

$$\begin{pmatrix} \mathbf{I}_2 & \mathbf{B} \\ \mathbf{0}_2 & \mathbf{A} \end{pmatrix}.$$

Now we use the symplectic property to show that $\mathbf{B}$ must be zero and $\mathbf{A}$ must be (2 by 2) symplectic. Indeed:

$$\widetilde{\mathbf{J}}(\mathbf{e}_1, \mathbf{e}_3) = 0 \Rightarrow \widetilde{\mathbf{J}}(\mathbf{e}_1,\ B_{11}\mathbf{e}_1 + B_{21}\mathbf{e}_2) = 0$$

as well. But this latter term is equal to B_{21} itself. Continuing in this way shows us that $\mathbf{B}$ is indeed zero. This clearly forces $\mathbf{A}$ to be symplectic in and of itself. But 2 by 2 symplectic matrices are the same as matrices of determinant 1, therefore we've proven that the double stabilizer

$$\left|\mathbf{Sp}_4\left(\mathbb{Z}_p\right)_{(\mathbf{e}_1,\mathbf{e}_2)}\right| = \left|\mathbf{SL}_2\left(\mathbb{Z}_p\right)\right| = (p+1) \cdot p \cdot (p-1).$$

Now for the more annoying part. Our last remaining obstacle is to figure what the size of the orbit of $\mathbf{e}_2$ is under the action of $\mathbf{Sp}_4\left(\mathbb{Z}_p\right)_{\mathbf{e}_1}$, i.e. to find $\#(\mathbf{Sp}_4\left(\mathbb{Z}_p\right)_{\mathbf{e}_1} \bullet \mathbf{e}_2)$. Since $\widetilde{\mathbf{J}}(\mathbf{e}_1, \mathbf{v})$ picks up the $\mathbf{e}_2$ component of $\mathbf{v}$ any element in the orbit of $\mathbf{e}_2$ must be of the form

$$\begin{bmatrix} a \\ 1 \\ b \\ c \end{bmatrix}.$$

[II] Recall that $|\,\mathcal{U}(\mathbb{Z}_n)\,|$ is, by definition, the Euler totient function, $\vartheta(n)$ first introduced back in 3.4.3.

[III] Alas, in this case we'll have more trouble calculating the size of the orbit, so we have the complete opposite problem!

Therefore the orbit can have at most p^3 elements in it. I claim that it has exactly this many, i.e. I claim that there's a symplectic transformation which fixes $\mathbf{e}_1$ and takes $\mathbf{e}_2$ to

$$\mathbf{v} := \begin{bmatrix} a \\ 1 \\ b \\ c \end{bmatrix}$$

for any a, b, and c in $\mathbb{Z}_p$. There are two cases:

- If $a \neq 0$, then the transformation, then we have that $\widetilde{\mathbf{J}}(\mathbf{e}_2, \mathbf{v}) \neq 0$, in fact it equals a, and so by part **(??)** of the exercise we've been referencing we know that

  $$t_{\mathbf{v}-\mathbf{e}_2}^{\lambda^{-1}}(\mathbf{e}_2) = \mathbf{v}.$$

  Moreover, since both $\mathbf{v}$ and $\mathbf{e}_2$ satisfy $\widetilde{\mathbf{J}}(*, \mathbf{e}_1) = 1$ the difference $\mathbf{v} - \mathbf{e}_2$ is $\mathbf{J}$ orthogonal to $\mathbf{e}_1$. Therefore the stated symplectic transformation fixes $\mathbf{e}_1$.
- If $a = 0$, then replace $\mathbf{v}$ with $t_{\mathbf{e}_1}(\mathbf{v}) := t_{\mathbf{e}_1}^1(\mathbf{v})$. Note that since $\widetilde{\mathbf{J}}(\mathbf{e}_1, \mathbf{e}_1) = 0$, this transformation does not move $\mathbf{e}_1$. Notice that

  $$t_{\mathbf{e}_1}(\mathbf{v}) = \begin{bmatrix} 1 \\ 1 \\ b \\ c \end{bmatrix}.$$

  Therefore we can proceed as before to move $\mathbf{e}_2$ to $t_{\mathbf{e}_1}(\mathbf{v})$ via $t_{t_{\mathbf{e}_1}(\mathbf{v})-\mathbf{e}_2}$ and the applying the inverse of symplectic transformation $t_{\mathbf{e}_1}$ to move the resulting vector finally to $\mathbf{v}$. Note that the net result fixes $\mathbf{e}_1$ since this is the case at every stage of our progress.

Therefore by **OST** we've proven that

$$\left|\mathbf{Sp}_4\left(\mathbb{Z}_p\right)_{\mathbf{e}_1}\right| = \#\left(\mathbf{Sp}_4\left(\mathbb{Z}_p\right)_{\mathbf{e}_1} \bullet \mathbf{e}_2\right) \cdot \left|\mathbf{Sp}_4\left(\mathbb{Z}_p\right)_{(\mathbf{e}_1,\mathbf{e}_2)}\right| = p^3 \cdot \left(p\left(p^2-1\right)\right).$$

Finally going back and plugging this result into Equation (18) finally gives us the formula we were looking for:

$$\left|\mathbf{Sp}_4\left(\mathbb{Z}_p\right)\right| = \left(p^4-1\right) \cdot \left|\mathbf{Sp}_4\left(\mathbb{Z}_p\right)_{\mathbf{e}_1}\right| = \left(p^4-1\right) \cdot p^3 \cdot \left(p\left(p^2-1\right)\right) = \boxed{p^4 \cdot \left(p^4-1\right) \cdot \left(p^2-1\right)}.$$

EXERCISE 10.58. Generalize the previous example and find a formula which computes the order of the group $\mathbf{Sp}_{2n}\left(\mathbb{Z}_p\right)$.

SOLUTION

EXERCISE Hard! 10.59. Prove that

$$\{t_{\mathbf{v}} : \mathbf{v} \in \mathbb{Z}_p^{2n}\}$$

generates $\mathbf{Sp}_{2n}\left(\mathbb{Z}_p\right)$.
Hint! You're almost there, but you have a few steps left to go.

SOLUTION

Exercise 10.60.

(1) Prove that for all $\mathbf{v} \in \mathbb{Z}_p^{2n}$ that $t_{\mathbf{v}}$ has determinant 1.

(2) Using the previous part and Exercise 10.59 to prove that every matrix in $\mathbf{Sp}_{2n}\left(\mathbb{Z}_p\right)$ has determinant 1.

Solution

10.9.3 Counting Cycle Types in $\mathcal{S}_n$

Recall by Proposition 10.100 that we have

$$\#\left(\mathrm{cc}_G(g)\right) \cdot |C_G(g)| = |G|. \tag{19}$$

relating them. In practice, this can be used in many ways. The most obvious 3 ways are

- To find the order of G by knowing the orders of both conjugacy class and centralizer of any element.
- To find the order of the centralizer by knowing the orders of the conjugacy class and the order of the groups.
- Of course, lastly to find the order of the conjugacy class if we know the orders of the group and the centralizer.

Let's look at the special case that $G = \mathcal{S}_n$. This is a case of the third type. This is because we already know the order of $\mathcal{S}_n$–$n!$– and it's fairly easy to find the order of the centralizer of an element in $\mathcal{S}_n$. For instance, in $\mathcal{S}_5$ there are 7 cycle types, namely the $\underline{1}$–cycles (the identity), $\underline{2}$–cycles, $\underline{3}$–cycles, $\underline{4}$–cycles, $\underline{5}$–cycles, $\underline{2}\,\underline{2}$–cycles, and $\underline{3}\,\underline{2}$–cycles [IV]. Now, if g centralizes a $\underline{2}$–cycle, say $x = (1\,2)(3)(4)(5)$, then since conjugation is replacement in $\mathcal{S}_n$ we have

$$gxg^{-1} = (g(1)\,g(2))(g(3))(g(4))(g(5))$$

and we must have that g permutes the 3 (implicit) fixed points of $(1\,2)$ and either fixes 1 and 2 or swaps them. There are $3! = 6$ ways to do this first thing and an additional 2 ways of doing this second. As such we have a total of 12 elements in the centralizer of $(1\,2) \in \mathcal{S}_5$. Therefore, by Proposition 10.100, there must be $120/12 = 10$ $\underline{2}$–cycles in $\mathcal{S}_5$. Since this number is small enough I will list them

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

This shouldn't be all that impressive in this case since it is clear that there are exactly as many $\underline{2}$–cycles as there are two element subsets of $\overline{n}$ which in general is calculated by the formula

$$\binom{n}{2} = \frac{n(n-1)}{2}.$$

More generally, for an $\underline{m}$–cycle, $x = (1\,2\,\cdots\,m) \in \mathcal{S}_n$ we must have that a g commuting with x is in the group $\langle x, \mathcal{S}_{\overline{n}^x} \rangle$ which has $m(n-m)!$ elements in it. Therefore, in $\mathcal{S}_n$ there are a total of

$$\frac{n!}{m(n-m)!}$$

$\underline{m}$–cycles (no, this isn't n choose m, there is no factorial after the m). This was originally given to you to prove in Exercise 7.1 back on page 347. In particular, there are $120/3 \cdot 2! = 20$ $\underline{3}$-cycles, $120/4 \cdot 1! = 30$ $\underline{4}$-cycles and $120/5 = 24$ $\underline{5}$-cycles in $\mathcal{S}_5$. For a $\underline{3}\,\underline{2}$-cycle, say $(1\,2\,3)(4\,5)$ it is still pretty basic to see that $g \in \langle (1\,2\,3), (4\,5) \rangle$

[IV] Of course $\underline{2}$-cycles in $\mathcal{S}_5$ are really $\underline{2}\,\underline{1}\,\underline{1}\,\underline{1}$-cycles

which has 6 elements in it (this group is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_3$. Therefore, there are a total of 20 $\underline{3}\,\underline{2}$–cycles in $\mathcal{S}_5$. It's a little trickier to figure out the centralizer of a $\underline{2}\,\underline{2}$–cycle, say $(1\,2)(3\,4)$ since we are allowed to swap the two $\underline{2}$–cycles. There are two ways we can swap these two to cycles as well as $2 \cdot 2 = 4$ ways we can maintain them. Therefore, the centralizer of a $\underline{2}\,\underline{2}$–cycle has 8 elements and therefore there are a total of $5!/8 = 5 \cdot 3 = 15$ $\underline{2}\,\underline{2}$–cycles in $\mathcal{S}_5$.

Let's try to come up with the general formula. For an element $\sigma \in \mathcal{S}_n$ whose proper cycle type has corresponding partition

$$(n^{i_n}, (n-1)^{i_{n-1}}, \ldots, 1^{i_1})$$

(so i_1 be the number of $\underline{1}$–cycles of σ ⟨i.e. the number of points fixed by σ ⟩, i_2 be the number of $\underline{2}$–cycles in the cycle decomposition of σ, etc). Now conjugation by an element which commutes with σ can only do a combination of fixing the i_m, $\underline{m}$–cycles of σ or swapping them amongst themselves. This contributes a $m^{i_m} \cdot i_m!$ factor to the size of the centralizer of σ (the power part counts the ways we can fix each $\underline{m}$–cycle and the $i_m!$ terms counts the ways we can permute the i_m $\underline{m}$–cycles). Putting these together we have that there are a total of

$$(n^{i_n}\, i_n!) \cdots (2^{i_2}\, i_2!) \cdot (1^{i_1} i_1!)^{\text{V}} \tag{20}$$

in the centralizer of σ. In conclusion:

THEOREM 10.113. *There are*

$$\frac{n!}{(n^{i_n}\, i_n!) \cdots (2^{i_2}\, i_2!) \cdot (1^{i_1} i_1!)} \tag{21}$$

elements in $\mathcal{S}_n$ with proper cycle type

$$(n^{i_n}, (n-1)^{i_{n-1}}, \ldots, 1^{i_1}).$$

EXAMPLE 10.114.

(1) As we already know there are

$$\frac{4!}{(2^2 \cdot 2!)} = 3$$

$\underline{2}\,\underline{2}$–cycles in $\mathcal{S}_4$ (here $\lambda = (2^2)$). In fact they are out old friends:

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

which are the 3 non-identity elements of the Klein four group.

(2) Let's figure out how many $\underline{2}\,\underline{2}\,\underline{2}$–cycles there are in $\mathcal{S}_6$. (So the partition in question is $\lambda = (2^3)$). Using the formula, we see that there are a total of

$$\frac{6!}{(2^3 \cdot 3!)} = \frac{\overset{3}{\cancel{6}} \cdot 5 \cdot \cancel{4}}{\cancel{2^3}} = 15$$

$\underline{2}\,\underline{2}\,\underline{2}$–cycles in $\mathcal{S}_6$. Notice that this happens to correspond with the number of $\underline{2} = (2, 1^4)$–cycles in $\mathcal{S}_6$,

$$\frac{6!}{(2^1 \cdot 1!)(1^4 \cdot 4!)} = \frac{6 \cdot 5}{2} = 15.$$

(This reflects that there is a highly non-trivial automorphism of $\mathcal{S}_6$ which exchanges all $\underline{2}\,\underline{2}\,\underline{2}$—cycles with $\underline{2}$–cycles.)

[V] Of course 1^{i_1} and $i_n!$ are both equal 1 (since we are in $\mathcal{S}_n$ we can only have 1 or 0 $\underline{n}$–cycles showing up in the cycle decomposition!) contribute nothing to the formula, but we write them in to make the general formula a little clearer.

(3) Let's figure out how many $\underline{3}\,\underline{3}\,\underline{2}$-cycles (so the partition is $\lambda = (3^2, 2, 1^2)$) there are in $\mathcal{S}_{10}$. For instance, $(1\ 3\ 6)(2\ 4\ 10)(5\ 9)$ is a perfectly good example of this. Here we implicitly have $10 - (3 + 3 + 2) = 2$ fixed points (so we are really looking at the partition $(3^2, 2, 1^2)$). Using the formula, we therefore have a total of

$$\frac{10!}{(3^2 \cdot 2!)(2^1 \cdot 1!)(1^2 \cdot 2!)} = 50,400$$

$\underline{3}\,\underline{3}\,\underline{2}$–cycles in $\mathcal{S}_{10}$.

(4) Let's figure out how many $\underline{4}\,\underline{4}\,\underline{4}\,\underline{4}\,\underline{3}\,\underline{3}\,\underline{2}$–cycles there are in $\mathcal{S}_{26}$.[VI] For starters, since $4 + 4 + 4 + 4 + 3 + 3 + 2 + 2 = 22$ and $26 - 22 = 4$, we have corresponding partition

$$\lambda = (4^4, 3^2, 2^2, 1^4).$$

Plugging this into the formula we see that we have

$$\frac{26!}{(4^4 \cdot 4!)(3^2 \cdot 2!)(2^2 \cdot 2!)(1^4 \cdot 4!)} = 25,324,030,189,458,000,000$$

or a (mere) 25 quintillion[VII] plus elements of $\mathcal{S}_{26}$ with this cycle type. This is only about 50 times larger than the estimated age of the universe in seconds. This number might sound huge, but it's nothing compared to the number

$$26! = 403,291,461,126,605,635,584,000,000.$$

Of course we could get really carried away and start talking about a deck of cards which has 52 "elements in it" so you can imagine how large the various numbers are in that group.

EXAMPLE 10.115. In Equation (5) on page 364 we showed that there are 15 different cycle types in $\mathcal{S}_7$. We record all the data we know about elements in these cycle types

[VI] Since $\pi(26) = 2436$, there are a total of 2436 cycle types in $\mathcal{S}_{26}$ ... nervous face :/

[VII] Yes! I admit I had to look up what the hell we call a number this big ... did you know that a number with 45 zeros after it is called a *quattuordecillion*? After hearing this word pronounced [here], I'm sure I've never heard this word before.

	Cycle Type	Example	Order	Parity	Number	
1	1	(1)	1	Even	1	$\mathbf{o}_1 = 1$
2	2	(1 2)	2	Odd	21	$\mathbf{o}_2 = 231$
3	2 2	(1 2)(3 4)	2	Even	105	
4	2 2 2	(1 2)(3 4)(5 6)	2	Odd	105	
5	3	(1 2 3)	3	Even	70	$\mathbf{o}_3 = 350$
6	3 3	(1 2 3)(4 5 6)	3	Even	280	
7	4	(1 2 3 4)	4	Odd	210	$\mathbf{o}_4 = 840$
8	4 2	(1 2 3 4)(5 6)	4	Even	630	
9	5	(1 2 3 4 5)	5	Even	504	$\mathbf{o}_5 = 504$
10	6	(1 2 3 4 5 6)	6	Odd	840	$\mathbf{o}_6 = 1470$
11	3 2	(1 2 3)(4 5)	6	Odd	420	
12	3 2 2	(1 2 3)(4 5)(6 7)	6	Even	210	
13	7	(1 2 3 4 5 6 7)	7	Even	720	$\mathbf{o}_7 = 720$
14	5 2	(1 2 3 4 5)(6 7)	10	Odd	504	$\mathbf{o}_{10} = 504$
15	4 3	(1 2 3 4)(5 6 7)	12	Odd	420	$\mathbf{o}_{12} = 420$

Using this table we can answer a question which reads as follows:

> An automatic card shuffler in a casino shuffles a deck of exactly 7 (distinct) cards the same way everytime. Not knowing how the machine works, what is the probability that if we let the machine operate 10 times that it returns the cards in exactly the same order as what was put into the machine?

Well the question is asking us to find the ratio

$$\frac{\mathbf{O}_{10}(\mathcal{S}_7)}{|\mathcal{S}_7|} = \frac{\mathbf{O}_{10}(\mathcal{S}_7)}{7!}.$$

But we know that

$$\begin{aligned}\mathbf{O}_{10}(\mathcal{S}_7) &= \sum_{d|10} \mathbf{o}_d(\mathcal{S}_7) \\ &= \mathbf{o}_1(\mathcal{S}_7) + \mathbf{o}_2(\mathcal{S}_7) + \mathbf{o}_5(\mathcal{S}_7) + \mathbf{o}_{10}(\mathcal{S}_7) \\ &= 1 + 231 + 504 + 504 \\ &= 1240.\end{aligned}$$

Therefore our answer is

$$\frac{1240}{7!} \approx 0.246$$

or just slightly less than 1/4.

EXERCISE 10.61. Make similar tables for the 11 cycle types of $\mathcal{S}_6$ and the 22 cycle types in $\mathcal{S}_8$. Use the first one it to calculate the number of elements of order 6 in $\mathcal{S}_8$ and the number of elements of order 2 in $\mathcal{S}_8$.

SOLUTION

MAPLE EXERCISE 10.62. Use a computer algebra system (or a fairly sophisticated calculator) to compute the probability of randomly shuffling a new, full deck of cards twice (in the same way) and ending up with an unshuffled deck. Compare this probability to the number seconds since the universe began (some 15 billion years ago).

SOLUTION

MAPLE EXAMPLE 10.116. Writing a little macro in MAPLE to compute the number of elements in a given conjugacy class seems to be more difficult that I would hope it would be. This is due to the fact that it seems MAPLE messed up the commands for this. For a partition involving $1's$, it seems to go nuts. This is the reason for the command titled tru below. It's only reason for being is to toss out the $1's$ at the end of a partition. So, for example,

$$\texttt{tru}([1,1,1,2,3,3,5]);$$
$$[2,3,3,5]$$

Here are the list of commands (as usual, the > signs are not to be copied into the code, they simply represent the individual chunks which should be copied over.)

```
> with(combinat):       #Loads package dealing w/partitions
> with(group):         #Loads group theory package
> fis:=proc(L)               #Calculates first index in input, L, neq 1
   local i;
   i:=1;
   while L[i]=1 do i:=i+1; od;
   return(i);
  end;
> tru:=proc(L)          #Uses fis to cut all 1's off input list, mentioned above
   local n,m, LL;
   m:=nops(L);
   n:=fis(L);
   LL:=L[n..m];
   return(LL);
  end;
> TT:=proc(n::integer)  #Defines command which returns standard n--cycle (123...n)
   local L, i;
   L:=[];
   for i from 1 to n do
```

```
    L:=[op(L),i];
    od;
    return([L]);
   end;
>  FFF:=proc(n)
    local i, m, L, LL, Sn;
    L:=map(tru,partition(n,n)[2..]);    #Would be easier if maple could load Sn directly
    Sn:=permgroup(n,{[[1,2]],TT(n)});  #Sn (w/gens (12) & (123 ... n))
    LL:=[ [[1],1], op(L)];
    m:=nops(L);
    for i from 1 to m do
    LL[i+1]:=[L[i],SnConjugates(Sn,L[i])];   #SnConjugates(Sn,L[i]) returns # of perms w/c.t. L[i]
    od;
    return(LL);
   end;
```

Now if we actually use our command FFF we will get a quick and easy way of computing the above table. For instance,

```
>  FFF(7);
```

```
[[[1],1],[[2],21],[[2,2],105],[[2,2,2],105],[[3],70],[[2,3],420],[[2,2,3],210],
[[3,3],280],[[4],210],[[2,4],630],[[3,4],420],[[5],504],[[2,5],504],[[6],840],[[7],720]]
```

which matches what we got above: for instance, the entry which entry in the output list which says $[[2,2,2],105]$ reflects the fact that there are 105 $\underline{2}\,\underline{2}\,\underline{2}$–cycles in $\mathcal{S}_7$.

EXERCISE 10.63.

(1) Show that an automorphism of $\mathcal{S}_n$ must take $\underline{2}$–cycles to $\underline{2}$–cycles except for the case that $n = 6$.
Hint! An automorphism of $\mathcal{S}_n$, like of any group, must take elements of a certain order to elements of that same order. Thus, an automorphism of $\mathcal{S}_n$ must take an element of order 2 to an element of order 2. However, the only elements of order 2 are products of disjoint $\underline{2}$–cycles:

$$(a_1 b_1)\cdots(a_m b_m).$$

For a particular m, show that the number of these is a multiple of

$$\binom{n}{2m}.$$

Therefore (why?) you only need to show that the number of $\underline{2}^{n/2}$–cycles is the same as the number of $\underline{2}$–cycles in $\mathcal{S}_n$ for all positive integer except for $n = 6$.

(2) Use Theorem 9.50 to conclude for all $n \neq 6$, $\mathrm{Aut}(\mathcal{S}_n) \cong \mathcal{S}_n$.

SOLUTION

10.9.4 Conjugacy Classes in $\mathcal{A}_n$*

How about in $\mathcal{A}_n$? What do the conjugacy classes look like there?

We start of by giving a general discussion. The formulas from Proposition 10.100 and Proposition 10.102 can be used at the same time to prove things. For instance, for $h \in H \leqslant G$ we have that

$$\#\mathfrak{cc}_H(h) = |H| / |C_H(h)|.$$

Like the stabilizer of a point under any action, we also have

$$C_H(h) = C_G(h) \cap H.$$

Using the mentioned formulas, we see that

$$\begin{aligned}
\#\mathfrak{cc}_H(h) &= \frac{|H|}{|C_H(h)|} \\
&= \frac{|H|}{|C_G(h) \cap H|} \\
&= \frac{|H \cdot C_G(h)|}{|C_G(h)|} \\
&= \frac{|H \cdot C_G(h)|}{|G|} \cdot \frac{|G|}{|C_G(h)|} \\
&= \frac{1}{[G : H \cdot C_G(h)]} \cdot \#\mathfrak{cc}_G(h).
\end{aligned}$$

or cross multiplying[VIII]

$$\#(\mathfrak{cc}_G(h)) = [G : H \cdot C_G(h)] \cdot \#(\mathfrak{cc}_H(h)).$$

For instance, if $H \leqslant G$ is maximal which happens if for instance H has prime index in G, then as a subgroup of G containing H, we have either $HC_G(h) = G$ if there exists some $g \in G \setminus H$ which commutes with h, in which case

$$\#\mathfrak{cc}_G(h) = \#\mathfrak{cc}_H(h).$$

Or $H \cdot C_G(h) = G$ if there does not exist some $g \in G \setminus H$ which commutes with h, in which case

$$\#\mathfrak{cc}_G(h) = \#\mathfrak{cc}_H(h) \cdot [G : H].$$

More generally $\#(H \cdot C_G(h))$ is equal to the number of cosets, gH, which can be represented by an element of $C_G(h)$. This is particularly easy to measure when we have a nice homomorphism with kernel H. For instance, if we take $G = \mathbf{GL}_2\left(\mathbb{Z}_p\right)$, $H = \mathbf{SL}_2\left(\mathbb{Z}_p\right)$, then

$$C_{\mathbf{GL}_2(\mathbb{Z}_p)}\left(\left(\begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix}\right)\right) = \left\{\begin{pmatrix} a & b \\ 0 & a \end{pmatrix} : a \in \mathbb{Z}_p^\times, b \in \mathbb{Z}_p\right\}.$$

Therefore the determinant of matrix which commutes with $\mathfrak{Q}$ has determinant a perfect square and as such $[G : H \cdot C_G(h)] = 2$. Therefore we have that $\mathfrak{Q}$ is conjugate to half as many matrices over $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ as it is over $\mathbf{GL}_2\left(\mathbb{Z}_p\right)$. For instance, one can see easily that the matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

[VIII] Actually, for a general action $G \circlearrowright X$, the same argument shows that

$$\#(G \bullet x) = [G : H \cdot G_x] \cdot \#(H \bullet x).$$

are **not** conjugate over $\mathbf{SL}_2(\mathbb{Z}_3)$, but they are conjugate over $\mathbf{GL}_2(\mathbb{Z}_3)$.

For example, if we take the $G = \mathcal{S}_n$ and $H = \mathcal{A}_n$, then we have for all even permutations σ that

$$\#\mathrm{cc}_{\mathcal{S}_n}(\sigma) = 2 \cdot \#\left(\mathrm{cc}_{\mathcal{A}_n}(\sigma)\right) \text{ or } \#\mathrm{cc}_{\mathcal{S}_n}(\sigma) = \#\mathrm{cc}_{\mathcal{A}_n}(\sigma)$$

depending on whether there is an odd permutation which commutes with σ or not.

Over $\mathcal{S}_n$ things are even easier. For example, it's not hard to see that the only even cycle types of permutations in $\preccurlyeq$ which do not commute with an even permutation are $\underline{7}\,\underline{1}$–cycles and $\underline{5}\,\underline{3}$–cycles. Therefore, any permutation of these cycle types are conjugate to $1/2$ as many permutations via $\mathcal{A}_n$ as they are via $\mathcal{S}_n$. However, for all other cycle types are conjugate to the same number of permutations.

In general

- **If σ contains a cycle of even length.** For if we call it τ, then τ is an odd permutation (cycle is even iff its length is odd) commuting with σ.
- **If σ fixes more than two points** Than the transposition swapping these two points is an odd permutation which commutes with σ since it is disjoint from σ.
- **If σ contains two (or more) cycles of the same length.** The first case already ruled out the possibility that this common length is even. Therefore, if we have two odd length cycles

  $$\tau_a = (a_1 \cdots a_{2k+1}) \quad \text{and} \quad \tau_b = (b_1 \cdots b_{2k+1})$$

  in σ, then the permutation

  $$\tau := (a_1\ b_1) \cdots (a_{2k+1}\ b_{2k+1})$$

  is odd, since the supports of τ_a and τ_b are disjoint and as such τ is the product of an odd number of transpositions and is therefore odd, and commutes with σ.

Indeed, these are the only types of even permutations which commute with an even permutation. What's left over? Well simply permutations with cycle type corresponding to a partition of n with

$$\lambda_1 > \cdots > \lambda_m \tag{22}$$

and each λ_i odd. Therefore, we are looking at sequences of distinct odd integers which add up to n.

Indeed, let σ be a permutation which satisfies (22). Then

$$\sigma = \tau_1 \cdots \tau_m$$

where each τ_i is a cycle of odd length and no two of the τ_i's have the same length (here we explicitly writing down any $\underline{1}$–cycles). As such each τ_i is an even permutation. Moreover, by our results about the centralizer of a general element of $\mathcal{S}_n$, the centralizer of σ is

$$\langle \tau_1, \ldots, \tau_m \rangle.$$

(This only depends on the fact that none of the lengths are the same.) Therefore, as a group generated by even permutations, we have

$$C_{\mathcal{S}_n}(\sigma) \subseteq \mathcal{A}_n.$$

Remark 10.117. Note we can see how many cycle types there are of this form by checking out the x^n coefficient in the polynomial

$$\prod_{i=0}^{\infty}(1 + x^{2i+1})$$

which (initially) does not grow too terribly quickly but it eventually does take off, see Figure 10.2. This is because the n^{th} coefficient of this polynomial measures exactly how many ways we can find distinct odd integers which add up to n.

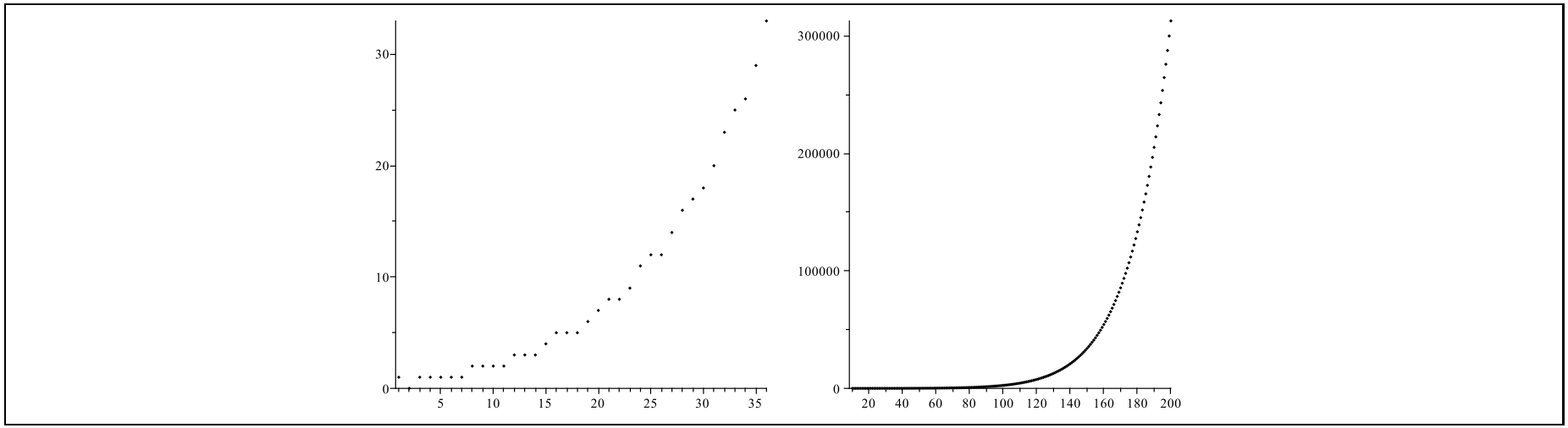


Figure 10.2: Number of even cycle types not commuting with any odd permutation, i.e. the number of conjugacy classes in $\mathcal{A}_n$ which have twice as many elements as they have in $\mathcal{S}_n$.

For instance, since there are a total of

$$\frac{8!}{5 \cdot 3} = 2688$$

$\underline{5}\,\underline{3}$–cycles in $\mathcal{S}_8$ and since a $\underline{5}\,\underline{3}$–cycle satisfies (22), we know that the set of $\underline{5}\,\underline{3}$–cycles (i.e. the conjugacy class of any given $\underline{5}\,\underline{3}$–cycle) split in half inside of $\mathcal{A}_n$. For example, the two $\underline{5}\,\underline{3}$–cycles

$$(1\,2\,3\,4\,5)(6\,7\,8) \quad \text{and} \quad (1\,2\,3\,4\,5)(6\,8\,7)$$

are no longer conjugate in $\mathcal{A}_8$. Therefore we have

$$\#\left(\mathrm{cc}_{\mathcal{A}_8}(1\,2\,3\,4\,5)(6\,7\,8)\right) = \frac{2688}{2} = 1344.$$

On the other hand, a $\underline{3}\,\underline{3}\,\underline{2}$–cycle does **not** satisfy (22) and so its conjugacy class remains intact when passing to $\mathcal{A}_n$ (indeed, the odd permutation $(7\,8)$ commutes with the $\underline{3}\,\underline{3}\,\underline{2}$–cycle $(1\,2\,3)(4\,5\,6)(7\,8)$). Therefore, in both $\mathcal{S}_8$ and $\mathcal{A}_8$ all

$$\frac{8!}{3^2 \cdot 2! \cdot 2} = 1120$$

$\underline{3}\,\underline{3}\,\underline{2}$–cycles live in a single conjugacy class.

10.9.5 Convex regular 4-polytopes*

In this optional section we talk about a particular 4 dimensional regular polytope, the 24–cell. I will not define what it means to be regular, but you should think of it as, like in 2 (regular polygons) and 3 dimensions (the 5 platonic solids), something which has a ton of symmetry. The cool thing is that, even though we can't really directly see this 4 dimensional thing, we will be able use group theory to at least get an impression of what it looks like. For instance, we will be able to know how many vertices (well, this much is by definition), edges, faces, and cells (three dimensional faces) the polytope has. For instance, with this object we have no hope of seeing directly, we will be able to show that it is made up of 96 triangles, and 24 icosahedrons inside of it (the 24 is the reason for its name, the 24 cell).

There are a total of 6 convex regular 4–polytopes living in $\mathbb{R}^4$[IX]. Like in every dimension, there is a (hyper) tetrahedron, a (hyper) cube, and a (hyper) octahedron. The vertices of the cube and octahedron as easy enough to guess: the cube having 16 vertices $(\pm 1, \pm 1, \pm 1, \pm 1)$, and the octahedron having 8 vertices $(\pm 1, 0, 0, 0)$, $(0, \pm 1, 0, 0)$,

[IX] It turns out that only in dimension 2 can you have an infinite number of polytopes

$(0,0,\pm 1,0)$, and $(0,0,0,\pm 1)$. The vertices of the tetrahedron are harder to give, but there are 5 of them; each at an equal distance from each other.

REMARK 10.118. It's typically easier to think of the vertices of an n–dimensional tetrahedron (called a n–simplex) as living in $n+1$ dimensions. There the vertices are simply stated as the $n+1$ standard basis vectors. This is the case even in dimension 2, where the regular 2-simplex is simply a regular triangle. To state the vertices of this triangle in the plain you need to deal with irrational numbers. But in $\mathbb{R}^3$ it can very simply be seen.

In 4 dimensions, we get 3 exceptional polytopes. They are known as the 24 cell,[X] the 120 cell, and the 600 cell. The numbers refer to the number of 3 dimensional faces of the objects. The latter two are dual to each other and former is self dual. The 24 cell has vertices given by $B\left(\blacktriangleright^3\right)$ and the 600 cell has vertices given by $B\left(\bigstar^3\right)$. (In many ways, 4 dimensions is much more like 2 dimensions than it is like 3 dimensions. Indeed, the regular polytopes in 2 dimensions had their vertices corresponding with the group $\{\{n\}\}$, the roots of unity. In 3 dimensions, all of the objects (like every object in the universe!) has a symmetry group, but it itself was not a symmetry group.

In three dimensions, the set of (o.p) symmetries were almost $\mathbb{S}^3$. In four dimensions, we have the set of (o.p) symmetries is almost the set of ordered pairs of quaternions.

THEOREM 10.119. *The following*

$$\begin{aligned}\star &: (\mathbb{S}^3 \times \mathbb{S}^3) \times \mathbb{R}^4 \to \mathbb{R}^4 \\ &: (q,p) \star \mathbf{V} = q\mathbf{V}p^{-1}.\end{aligned}$$

defines an action of $\mathbb{S}^3 \times \mathbb{S}^3$ by orthogonal transformations of positive determinant (i.e. elements of $\mathbf{SO}_4(\mathbb{R})$) whose permutation representation

$$: \mathbb{S}^3 \times \mathbb{S}^3 \to \mathbf{SO}_4(\mathbb{R}) : (q,p) \mapsto (q,p)\star$$

is surjective and has kernel consisting (only) of the two elements $\{(1,1),(-1,-1)\} \subseteq \mathbb{S}^3 \times \mathbb{S}^3$.

PROOF.

(1) Next we verify that the action is by orthogonal transformations. That is we need to check that

(a) ($(q,p) \star$ **is linear**) Let $\mathbf{V}$ and $\mathbf{W}$ be 4-vectors and $c \in \mathbb{R}$. Then

$$\begin{aligned}(q,p) \star (\mathbf{V} + c\mathbf{W}) &= q(\mathbf{V} + c\mathbf{W})p^{-1} \\ &= [q(\mathbf{V} + c\mathbf{W})]p^{-1} \\ &= (q\mathbf{V} + cq\mathbf{W})p^{-1} \\ &= q\mathbf{V}p^{-1} + cq\mathbf{W}p^{-1} \\ &= (q,p) \star \mathbf{V} + c(q,p) \star \mathbf{W}.\end{aligned}$$

(b) ($(q,p) \star$ **preserves lengths**) Thinking of $\mathbb{R}^4$ as $\mathbb{H}$, recall that for all $\mathbf{V} \in \mathbb{R}^4$, $\|\mathbf{V}\|^2 = \mathbf{V}\overline{\mathbf{V}}$. Then for any

$$(q,p) \in \mathbb{S}^3 \times \mathbb{S}^3 \text{ and } \mathbf{V} \in \mathbb{R}^4$$

[X] A fancier name for the 24 cell is the *icositetrachoron*.

we have

$$\begin{aligned}
\|(q,p)\star\mathbf{V}\|^2 &= ((q,p)\star\mathbf{V})*\overline{((q,p)\star\mathbf{V})}\\
&= \left(q\mathbf{V}p^{-1}\right)*\overline{(q\mathbf{V}p^{-1})}\\
&= \left(q\mathbf{V}p^{-1}\right)*\left(p\overline{\mathbf{V}}q^{-1}\right)\\
&= q\mathbf{V}\left(p^{-1}\overset{1}{\cancel{p}}\right)\overline{\mathbf{V}}q^{-1} && \text{(Reassociate)}\\
&= q\left(\mathbf{V}\overline{\mathbf{V}}\right)q^{-1}\\
&= q\,\|\mathbf{V}\|^2\,q^{-1}\\
&= \overset{1}{\cancel{qq^{-1}}}\,\|\mathbf{V}\|^2 && (\|\mathbf{V}\|\in\mathbb{R}\Rightarrow\|\mathbf{V}\|\in\mathcal{C}(\mathbb{H}^\times))\\
&= \|\mathbf{V}\|^2.
\end{aligned}$$

(c) $((q,p)\star$ **has determinant** $+1)$

Step 1 **(Quaternions of the form** (q,q)**, i.e. 3d-rotations)** Recall from our discussion of 3 dimensions (found on page 779), for all $q\in\mathbb{S}^3$, and $\mathbf{v}\in\mathbb{R}^3$, $(q,q)\star\mathbf{v}=q\mathbf{v}q^{-1}$ is a rotation (of $\mathbb{R}^3$) and so has determinant $+1$. Since $(q,q)\star 1=1$, we have that the (four by four!) matrix of $(q,q)\star$ has the form

$$\begin{pmatrix}1&0&0&0\\0&r_{22}&r_{23}&r_{24}\\0&r_{32}&r_{33}&r_{34}\\0&r_{42}&r_{43}&r_{44}\end{pmatrix}$$

with the 3 by 3 submatrix

$$\begin{pmatrix}r_{22}&r_{23}&r_{24}\\r_{32}&r_{33}&r_{34}\\r_{42}&r_{43}&r_{44}\end{pmatrix}$$

a 3 by 3 rotation matrix (and hence has determinant 1). This implies that $(q,q)\star$ has determinant 1.

Step 2 **(Quaternions of the form** $(q,1)$**, i.e. roto-translations)** Meanwhile for any $q\in\mathbb{S}^3$ we have that

$$(q,1)\star\mathbf{V}=q*\mathbf{V}*1^{-1}=q*\mathbf{V}^{\text{XI}}$$

for any $\mathbf{V}\in\mathbb{R}^4$. Since it corresponds to multiplication, and since $\mathbb{H}^\times$ is a group, $(q,1)\star$ fixes only the zero vector $\mathbf{0}\in\mathbb{R}^4$ (unless $q=1$ then it fixes everything). As a symmetry of $\mathbb{R}^4$ fixing only the origin, $(q,1)\star$ is the product of 4 reflections and, as such, has determinant $(-1)^4=1$.

Step 3 **(General case)** We write the general case as the composition of the first two cases. Let $(q_1,q_2)\in\mathbb{S}^3\times\mathbb{S}^3$. Then since $\mathbb{S}^3\preccurlyeq\mathbb{H}^\times$, $q_1q_2^{-1}=p$ for some

[XI] Here, as before, $*$ denotes the quaternion multiplication for emphasis.

$p \in \mathbb{S}^3$. So

$$\begin{aligned}(q_1, q_2) &= (q_1(q_2 q_2^{-1}), q_2)\\ &\overset{p}{=} ((q_1 q_2^{-1}) * q_2, q_2)\\ &= (pq_2,\ q_2)\\ &= (p * q_2,\ 1 * q_2)\\ &= (p,\ 1)(q_2,\ q_2) \qquad \text{(Def'n of mult on direct products)}\\ \Rightarrow (q_1,\ q_2)\star &= ((p,\ 1)\star) \circ ((q_2, q_2)\star).\end{aligned}$$

But we've already shown in **Step 1** that $(q_2, q_2)\star$, as a 3d-rotation, has determinant 1, and then in **Step 2** we also showed that $(p, 1)\star$, as a roto-translation, has determinant 1 as well. Therefore, since linear transformations of determinant 1 are closed under composition, their composition $(q_1, q_2)\star$ also has determinant 1.

(2) If (q, p) is in the kernel of the action, then

$$\forall \mathbf{V} \in \mathbb{R}^4,\ q\mathbf{V}p^{-1} = \mathbf{V} \tag{23}$$

for all four vectors $\mathbf{V} \in \mathbb{R}^4$ (i.e. quaternions). In particular, this is the case for the

$$\mathbf{V} = 1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}.$$

Therefore, we have

$$q * 1 * p^{-1} = 1 \Rightarrow q = p \Rightarrow (q, p) = (q, q).$$

Plugging this restriction into (23), we now know that we have

$$\forall \mathbf{V} \in \mathbb{R}^4,\ q\mathbf{V}q^{-1} = \mathbf{V}. \tag{24}$$

Identifying $\mathbb{R}^4$ with $\mathbb{H}$, this means that q (which is given to be an element of $\mathbb{S}^3$) commutes with every $\mathbf{V} \in \mathbb{H}$, in particular with every element of $\mathbb{S}^3$, i.e. $q \in \mathcal{C}(\mathbb{S}^3)$ (the center of $\mathbb{S}^3$). But we know that the center of $\mathbb{S}^3$ consists only of $\{\pm 1\}$. In conclusion,

$$(q, p) = (1, 1) \text{ or } (q, p) = (-1, -1)$$

as promised.

(3) It only remains to show that $\mathbb{S}^3 \times \mathbb{S}^3 \to \mathbf{SO}_4(\mathbb{R})$ is surjective. Let $\rho \in \mathbf{SO}_4(\mathbb{R})$. Then since $(\rho(1), 1)\mathbb{S}^3 \times \mathbb{S}^3$ does the same thing to $1 \in \mathbb{R}^4$ (again, identifying $\mathbb{R}^4$ with $\mathbb{H}$) as ρ does, we have that

$$\rho' := ((\rho(1))^{-1}, 1) \circ \rho$$

is (as the composition of elements of $\mathbf{SO}_4(\mathbb{R})$) an element of $\mathbf{SO}_4(\mathbb{R})$ which fixes 1. Therefore, as an element of $\mathbf{SO}_4(\mathbb{R})$ fixing 1, ρ' also fixes the set of all quaternions perpendicular with 1, which happen to be the set of all pure quaternions, which as usual, are identified with $\mathbb{R}^3$. This means we can think of ρ' as an element of $\mathbf{SO}_3(\mathbb{R})$. But since we've seen in the appendix, the map $\mathbb{S}^3 \to \mathbf{SO}_3(\mathbb{R})$ is surjective, there exists a $q \in \mathbb{S}^3$ so that for all $\mathbf{v} \in \mathbb{R}^3$

$$q\mathbf{v}q^{-1} = \rho'(\mathbf{v}).$$

But since $(\mathsf{q},\mathsf{q})\star\mathbf{v}=\mathsf{q}\mathbf{v}\mathsf{q}^{-1}$, we have that

$$(\mathsf{q},\mathsf{q})\star=\rho'\Rightarrow(\mathsf{q},\mathsf{q})\star=((\rho(1))^{-1},1)\circ\rho\Rightarrow(\rho(1)\mathsf{q},\mathsf{q})=(\rho(1),1)\circ(\mathsf{q},\mathsf{q})=\rho.$$

Therefore we have proved that under the map $\mathbb{S}^3\times\mathbb{S}^3\to\mathbf{SO}_4(\mathbb{R})$ that $(\rho(1)\mathsf{q},\mathsf{q})\mapsto\rho$. Since ρ is surjective we have completed the proof. ∎

Exercise 10.64. Prove that $(\mathsf{q},\mathsf{p})\in\mathbb{S}^3\times\mathbb{S}^3$ has a non-trivial (i.e. contains more that $\mathbf{0}\in\mathbb{R}^4$) fixed point set iff the real parts of q and p are the same.
Hint! First prove that the hypothesis is equivalent to the requirement that q and p are conjugate elements of $\mathbb{S}^3$. Then use Example/Theorem 5.102.

SOLUTION

We denote the 24 cell as C_{24} or C and its symmetry group by G and its corresponding quaternion group, i.e. the subgroup of $\mathbb{S}^3\times\mathbb{S}^3$ preserving C_{24} as BG. By definition, C_{24} has vertices $B\left(\blacktriangleright^3\right)$. Let's figure out how many symmetries it has by applying the OST. Indeed, the stabilizer of $1\in C_{24}$ under the action of BG, BG_1, consists of the the set of of pairs of quaternions $(\mathsf{q},\mathsf{q}')\in BG$ with

$$\mathsf{q}\mathsf{q}'^{-1}=\mathsf{q}1\mathsf{q}'^{-1}=(\mathsf{q},\mathsf{q}')\star 1=1\Rightarrow\mathsf{q}=\mathsf{q}'.$$

Moreover, we also need that (q,q') is an element of BG in the first place. That is we still need (q,q) to permute the other 23 vertices of C_{24} amongst themselves. That is for all $(\mathsf{q},\mathsf{q})\in BG_1$ and $Q\in B\left(\blacktriangleright^3\right)$ (i.e. a vertex of C_{24}), we have that $(\mathsf{q},\mathsf{q})\star Q=\mathsf{q}Q\mathsf{q}^{-1}$ is still an element of $B\left(\blacktriangleright^3\right)$ (i.e. still a vertex of C_{24}). That is the stabilizer of $1\in C_{24}$ under the action is

$$\left\{(\mathsf{q},\mathsf{q}):\mathsf{q}\in\mathrm{N}_{\mathbb{S}^3}\left(B\left(\blacktriangleright^3\right)\right)\right\}\cong\mathrm{N}_{\mathbb{S}^3}\left(B\left(\blacktriangleright^3\right)\right).$$

where $\mathrm{N}_{\mathbb{S}^3}\left(B\left(\blacktriangleright^3\right)\right)$, as for any group, denotes the normalizer of $B\left(\blacktriangleright^3\right)$ in $\mathbb{S}^3$, i.e.

$$\mathrm{N}_{\mathbb{S}^3}\left(B\left(\blacktriangleright^3\right)\right)=\left\{\mathsf{q}\in\mathbb{S}^3:\mathsf{q}B\left(\blacktriangleright^3\right)\mathsf{q}^{-1}=B\left(\blacktriangleright^3\right)\right\}.$$

Well, first we note that Q_8 is the unique subgroup of $B\left(\blacktriangleright^3\right)$ of order 8.[XII] Moreover, since the map $Q\mapsto\mathsf{q}Q\mathsf{q}^{-1}$ is linear, any quaternion normalizing Q_8 must also normalize $B\left(\blacktriangleright^3\right)$. Thus

$$\mathrm{N}_{\mathbb{S}^3}\left(B\left(\blacktriangleright^3\right)\right)=\mathrm{N}_{\mathbb{S}^3}\left(Q_8\right).$$

But the 6 elements of Q_8, $\{\pm\mathbf{i},\pm\mathbf{j},\pm\mathbf{k}\}$, are also preserved by the normalizer of Q_8. Again, a quaternion which permutes these 6 quaternions (via conjugation) uniquely determine an element of the normalizer of all of $B\left(\blacktriangleright^3\right)$. But, these 6 vertices are the vertices of an octahedron. Thus an element of $\mathbb{S}^3$ normalizes $\{\pm\mathbf{i},\pm\mathbf{j},\pm\mathbf{k}\}$ iff it permutes the vertices of an octahedron, (under the action of $\mathsf{q}\star\mathbf{v}=\mathsf{q}\mathbf{v}\mathsf{q}^{-1}$ considered in the 3 dimensional case) i.e. iff it's an element of $B\left(\blacklozenge^3\right)$ (the double of the symmetries of a cube =symmetries of an icosahedron). In particular, the stabilizer of 1 in BG has $\left|B\left(\blacklozenge^3\right)\right|=48$ elements.

Next, what's the orbit of $1\in C_{24}$ under the action of BG? Well, for any quaternion $\mathsf{q}\in B\left(\blacktriangleright^3\right)$ we have that $(\mathsf{q},1)\star\mathbf{v}=\mathsf{q}*\mathbf{v}$, the quaternion product of q and $\mathbf{v}\in\mathbb{R}^4$, and preserves C_{24}, since it preserves its vertices by closure of the group $B\left(\blacktriangleright^3\right)$ by multiplication. Thus the action is transitive and so the orbit of 1 consists of all 24 elements of $B\left(\blacktriangleright^3\right)$.[XIII]

Now we are double counting since the action has a kernel (consisting of two elements), but we are not counting any orientation reversing symmetries. Thus we need to divide our total $48\cdot 24$ by 2 and then multiply it by 2, hence we've done nothing to our total and we've proven

[XII] We will see this immediately in the next chapter as a consequence of the Sylow's theorems. For now you should notice that Q_8 consists of each every one of the elements whose order divides 8.

[XIII] We have actually shown that every element of BG can be written uniquely as $(\mathsf{q}\mathsf{q}',\mathsf{q}')$ where $\mathsf{q}\in B\left(\blacktriangleright^3\right)$ and $\mathsf{q}'\in B\left(\blacklozenge^3\right)$.

There are $48 \cdot 24 = 1152$ total symmetries of the 24 cell and $\frac{48 \cdot 24}{2} = 576$ orientation preserving ones.

Next, how many edges does this 24 cell have? Two distinct vertices, like in any polytope, in $B\left(\blacktriangleright^3\right)$ are connected by an edge iff their distance is as small as possible. This is clearly the same as the requirement that the dot product be as large as possible. Indeed, all vertices of (elements of) $B\left(\blacktriangleright^3\right)$ are in $\mathbb{S}^3$, and as such, have length 1. Therefore the distance between any two such vertices is

$$d^2(q_1, q_2) = (q_1 - q_2) \bullet (q_1 - q_2) = 1 + 1 - 2q_1 \bullet q_2 = 2(1 - q_1 \bullet q_2).$$

For instance the 8 vertices which are connected with the vertex 1 are the 8 vertices in $B\left(\blacktriangleright^3\right)$ with real part 1/2.

$$\begin{array}{ll} \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k} & \frac{1}{2} - \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k} \\ \frac{1}{2} + \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k} & \frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k} \\ \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k} & \frac{1}{2} - \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k} \\ \frac{1}{2} + \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k} & \frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k}. \end{array}$$

Notice that these are the 8 elements of order 6 in both of the groups $B\left(\blacklozenge^3\right)$ and $B\left(\blacktriangleright^3\right)$.

Exercise 10.65. Prove for quaternions q and q' in $\mathbb{S}^3$ of finite order (hence all points in $B\left(\blacktriangleright^3\right)$) have the same distance from 1 iff they have the same order.
Hint! As pointed out this is the same as requiring that both have the same (4 dimensional) dot product with 1. But the dot product of a quaternion with 1 is the pure part of that quaternion. Lastly a unit quaternion's order is determined entirely by its order.

SOLUTION

Let's look at the stabilizer of the edge denoted $[1, \omega]$ with endpoints 1 and ω, which denotes, as usual, the quaternion

$$\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}$$

(the first one listed of the 8 adjacent quaternions to 1 about).

I claim that the stablizer of the edge $e := [1, \omega]$ under the action of $BG \circlearrowright \text{Edges}(C_{24})$ is the set of quaternions

$$\{(q_1 q_2, q_2) : q_1 = 1 \text{ and } q_2 \omega q_2^{-1} = \omega\} \cup \{(q_1 q_2, q_2) : q_1 = \omega \text{ and } q_2 \omega q_2^{-1} = \omega^{-1}\}.$$

We denote the first part of this as BG_e^1 and the second part as BG_e^2. (This union is in bijection with the normalizer of $\langle \omega \rangle$.) Notice that if $(q_2, q_2) \in BG_e^1$, then q_2 is in the centralizer of ω and

$$(q_2, q_2) \star 1, \omega = [q_2 q_2^{-1}, q_2 \omega q_2^{-1}] = [1, \omega]$$

and it fixes both endpoints of the edge e. On the other hand, if $(\omega q_1, q_1) \in BG_e^2$, then

$$\begin{aligned} (\omega q_1, q_1) \star 1, \omega &= [\omega q_1 q_1^{-1}, \omega q_1 \omega q_1^{-1}] \\ &= [\omega, \omega\omega^{-1}] \\ &= [\omega, 1] \end{aligned}$$

and BG_e^2 exchanges the two endpoints the edge.

Now, as already pointed out, we can factor every symmetry of C_{24} as $(q_1q_2, q_2) = (q_1, 1)(q_2, q_2)$ with $q_1 \in B\left(\blacktriangleright^3\right)$ and $q_2 \in B\left(\blacklozenge^3\right)$. For a symmetry to preserve the edge $[1, \omega]$ it must either fix both 1 and ω or it must swap the two endpoints. In the first case, we must have that $q_1 = 1$ since this determines where 1 goes, so

$$(q_1q_2, q_2) = (q_2, q_2) \text{ and } (q_2, q_2) \star \omega = \omega \Rightarrow q_2\omega q_2^{-1} = \omega \Rightarrow q_2\omega = \omega q_2.$$

That is we must have that q_2 must commute with ω. Therefore, in this case, we have that $(q_1q_2, q_2) \in BG_e^1$.

What does it mean for a pair (q_1q_2, q_2) ($q_1 \in B\left(\blacktriangleright^3\right)$ and $q_2 \in B\left(\blacklozenge^3\right)$) to flip ω and 1? Well, since the pair (q_1q_2, q_2) takes 1 to ω we have $(q_1, 1)(q_2, q_2) \star 1 = \omega$ which implies $q_1 = \omega$. Now using the fact that this same element takes ω back 1, i.e., $(q_1q_2, q_2) \star \omega = 1$, we see

$$(\omega, 1)(q_2, q_2) \star \omega = 1 \Rightarrow \omega q_2\omega q_2^{-1} = 1 \Rightarrow q_2\omega q_2^{-1} = \omega^{-1}. \tag{25}$$

In this second case we now see that our starting guy, (q_1q_2, q_2) is an element of BG_e^2. Therefore we have proven that $BG_e = BG_e^1 \cup BG_e^2$. Now how many elements are in there?

Since $\mathcal{U}(\mathbb{Z}_6)$ contains only two elements, namely 1 and $5 = -1$, we have that the function

$$(q_1q_2, q_2) \mapsto q_2 \in B\left(\blacklozenge^3\right)$$

is an isomorphism from BG_e to the normalizer of ω in $B\left(\blacklozenge^3\right)$. Therefore, in total, the set of elements which fix the edge $[1, \omega]$, i.e. the elements of BG_e, has the same order as the normalizer of $\langle\, \omega \,\rangle \subseteq B\left(\blacklozenge^3\right)$. But this is easily seen to have 12 elements and be generated by ω and

$$\frac{\sqrt{2}}{2}(\mathbf{i} - \mathbf{j}).^{\text{XIV}}$$

This is because the centralizer of ω consists of 6 elements. Therefore BG_e^1 also has 6 elements. Moreover, we clearly have that for all $\mathbf{g}_1, \mathbf{g}_2 \in BG_e^2$, $\mathbf{g}_1\mathbf{g}_2 \in BG_e^1$. This is because both $\mathbf{g}_1$ and $\mathbf{g}_2$ exchange the vertices of $e = [1, \omega]$ by assumption, so their product must leave both vertices alone. Therefore BG_e^1 has index 2 in BG_e and so BG_e has order twice that of BG_e^1, i.e. it has order 12.

Another way to see that this normalizer, and hence BG_e, has 12 elements in it is to note that there are a total of 8 elements of order 6 in $B\left(\blacklozenge^3\right)$ (including ω) which consist of 4 pairs of such elements and their inverses. Thus there are a total of 4 subgroups of $B\left(\blacklozenge^3\right)$ of order 6. Since all 4 subgroups of order 6 are conjugate in $B\left(\blacklozenge^3\right)$, indeed they are conjugate even in $B\left(\blacktriangleright^3\right)$, the normalizer of a particular one has order $\left|B\left(\blacklozenge^3\right)\right|/4 = 48/4 = 12$. Since all 8 elements of order 6 are conjugate in $B\left(\blacktriangleright^3\right)$ the stabilizer of 1 acts transitively on the set of edges originating from 1. Using this and the fact that the symmetries of C_{24} act transitively on the set of its vertices, we can easily see that the symmetries of C_{24} also act transitively on the set of its edges.

The polytope C_{24} has

$$\frac{1152}{12} = 48 \cdot 2 = 96$$

edges.

Now what does it mean for three vertices to span a (2 dimensional) face? Well, the faces are simply the triples, $\{q_1, q_2, q_3\}$ with each pair connected by an edge

$$q_1 \bullet q_2 = q_2 \bullet q_3 = q_1 \bullet q_3 = \frac{1}{2}.$$

[XIV] This is a guy who conjugates ω to ω^{-1}. In fact any of the 6 pure quaternions perpindicular to the pure part of $\omega \sim i + j + k$ would do the same thing.

For instance there are 3 faces containing the vertices 1 and ω. They have the third vertex given by

$$\frac{1}{2} - \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}$$
$$\frac{1}{2} + \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}$$
$$\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} - \frac{1}{2}\mathbf{k}$$

The 8 vertices connected to 1 clearly form a 3–dimensional cube inside of C_{24}. Two vertices in this cube are joined by an edge in this cube iff along with the vertex 1, they span a (2–dimensional triangular) face. Thus 1 is contained inside of 12 faces. Since the faces are triangles there are 3 edges per face. On the other hand we just showed that there are a total of 3 faces per edge. Hence there are as many triangular faces of C_{24} as there are edges. Therefore we see that

The polytope C_{24} has

96

triangular faces.

Next, what are the three dimensional facets of C_{24}?. Well, as we have been doing, since the action is transitive on vertices, we look around the origin 1 for our hints. Since the 8 vertices adjacent to 1 forms a cube, which does not contain any triangles, the three dimensional faces of C_{24} cannot be tetrahedra. Indeed, the smallest loop in these 8 vertices, like in the cube, consists of squares. For instance the four vertices

$$\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k},\ \frac{1}{2} - \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}\ \frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}\ \frac{1}{2} + \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}$$

form the vertices of such a square. Notice the vertex, i.e. the element $\mathbf{k} \in B\left(\blacktriangleright^3\right) = \mathrm{Vert}(C_{24})$, has dot product equal to $1/2$ with each of these four vertices, it is adjacent to all four. This works out well since 1 is perpendicular to $\mathbf{k}$. Thus the vertices

$$\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k},\ \frac{1}{2} - \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}\ \frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}\ \frac{1}{2} + \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}\ 1,\ \mathbf{k}$$

give us 6 vertices of an octahedron inside of C_{24}. Moreover, for each of the six (square) faces of the cube inside of the 8 vertices adjacent to 1 there exists a unique basic quaternion, i.e. element of $\pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}$ so that the vertices in the square, the vertex 1, and the basic quaternion span an octahedron inside of C_{24}. In particular, every vertex is contained inside of 6 cells, which as octahedrons, each of which have 6 vertices. Therefore there are as many cells (3–dimensional faces) inside of C_{24} as there are vertices. That is

The polytope C_{24} has

24

cells. (This of course is how it got its name, the 24 cell!)

Remark 10.120. Depending on how much you wanted to cheat, you could have used several facts which would have made your life easier. For instance, it might appear like a coincidence that the number of vertices and cells matched, while at the same time the number of edges and faces also matched. However this is a direct consequence of the fact that the 24 cell is self dual. However, since I can't directly see this polytope, I see no fast way to of directly seeing that this dude is self dual!.

Exercise 10.66. Wiki the 24 cell,

http://en.wikipedia.org/wiki/24-cell

and check out the cool pics.

Solution

10.9.6 More Applications of OST

Notice for a finite group G, every element that is not in the center has more than one element in its conjugacy class. In fact this classifies non-central elements. That is there are two types of conjugacy classes in G: ones that contain exactly one element, i.e. conjugacy classes of central elements, and ones that contain more than one element, i.e. conjugacy classes on non-central elements. In the theorem below $\mathbf{K}$ is a subset of G which contains one and only one representative from the conjugacy class of every non-central element.

10.9.6.1 The Class Equation

The following theorem is used extensively in many other textbooks on abstract algebra. However, here we will try to limit its use. We present it more for completeness than for its uses. It should be clear that it follows almost directly from the **OST**.

Theorem 10.121 (The Class Equation). *For any finite group G*

$$|G| = |\mathcal{C}(G)| + \sum_{x \in \mathbf{K}} \frac{|G|}{|C_G(x)|}.$$

Proof. As pointed out already every element of G is conjugate to exactly one element of $\mathcal{C}(G) \coprod \mathbf{K}$. Since conjugacy classes partition G, we must have that

$$\begin{aligned} |G| &= \sum_{x \in \mathcal{C}(G) \cup \mathbf{K}} \#\mathrm{cc}_G(x) \\ &= \sum_{x \in \mathcal{C}(G)} \#\mathrm{cc}_G(x) + \sum_{x \in \mathbf{K}} \#\mathrm{cc}_G(x) \\ &= \sum_{x \in \mathcal{C}(G)} 1 + \sum_{x \in \mathbf{K}} \#\mathrm{cc}_G(x) \qquad ((x \in \mathcal{C}(G) \Rightarrow \mathrm{cc}_G(x) = \{x\} \Rightarrow \#\mathrm{cc}_G(x) = 1)) \\ &= |\mathcal{C}(G)| + \sum_{x \in \mathbf{K}} \#\mathrm{cc}_G(x). \end{aligned} \tag{26}$$

But we also have that

$$\#\mathrm{cc}_G(x) = \frac{|G|}{|C_G(x)|},$$

by Proposition 10.100 (which is simply **OST** for the action of G on itself by conjugation). Therefore, plugging this into (26), the sum becomes

$$|G| = |\mathcal{C}(G)| + \sum_{x \in \mathbf{K}} \frac{|G|}{|C_G(x)|}$$

as desired. ∎

Let's generalize this a bit. For any group action, there are two types of orbits orbits of points who are in global fixed point set of the action, and those who are not. Points in the global fixed point set of the action are classified by the fact that their orbits consist only of a single point, namely the original point itself. Now if

$$\{x_1, \ldots, x_n, y_1, \ldots, y_m\}$$

represent all orbits of $G \circlearrowright X$, where we have each $x_i \in X^P$ and each $y_i \notin X^P$, then we have $G \bullet \{x_i\} = \{x_i\}$. Moreover since every element of the global fixed point has an orbit containing only itself, we must have

$$\{x_1, \ldots, x_n\} = X^G.$$

Putting these two facts together we see

$$\begin{aligned} X &= G \bullet x_1 \bigsqcup \cdots \bigsqcup G \bullet x_n \bigsqcup G \bullet y_1 \amalg \cdots \amalg G \bullet y_m \\ &= \{x_1\} \bigsqcup \cdots \bigsqcup \{x_n\} \bigsqcup (G \bullet y_1 \amalg \cdots \amalg G \bullet y_m) \\ &= \{x_1, \ldots, x_n\} \bigsqcup (G \bullet y_1 \amalg \cdots \amalg G \bullet y_m) \\ &= X^G \bigsqcup (G \bullet y_1 \amalg \cdots \amalg G \bullet y_m). \end{aligned}$$

Taking orders shows us that

$$\#X = \#X^G + \sum \#(G \bullet y) \tag{27}$$

$$= \#X^G + \sum \frac{|G|}{|G_y|} \qquad \text{(By OST)} \tag{28}$$

where the sums are over a set of representatives of orbits of non-trivial elements of $G \circlearrowright X$.

Recall that many groups are centerless, i.e. have only the identity element which commute with every element of the group. Surprisingly, groups whose order is a positive power of a prime integer are never centerless. Such groups have a special name:

Definition 10.122. Let p be a prime integer. A group is a p-group if its order is a positive power of p. We call a group a $\mathfrak{p}$-group if it is a p-group for some prime number p.

Exercise 10.67. For a prime p, prove that every subgroup and the image of any homomorphism of a p–group is a p–group.

SOLUTION

Therefore, groups of order 5 or 25 are 5-groups; groups of orders 4, 32, or 256 are 2-groups; groups of order 49 or 96,889,010,407 are 7-groups, etc. Of course, any group of any of these orders are $\mathfrak{p}$-groups. A quick glance at the table in Appendix I shows that, up to isomorphism, most groups are $\mathfrak{p}$–groups.

In many cases the actual numbers present in the class equation are not that important. The proof of the following theorem illustrates this.

Theorem 10.123. *Suppose that G is a $\mathfrak{p}$–group. Then G is not centerless.*

Proof. By the class equation, we have the order of G, $|G|$, is written as 1 plus a sum of divisors of $|G|$ besides 1. Therefore, for instance since $|G| = p^r$ for some positive integer r and prime p, then every divisor of $|G|$ which is not the number 1, is also of the form p^s for some $1 \leq s \leq r$. Hence if such a group G were centerless, we'd have

$$p^r = 1 + p^{s_1} + \cdots + p^{s_m}$$

for integers $1 < s_i \leq r$. But taking this equation mod p shows that $[\,0\,]_p = [\,1\,]_p$, which is false. Thus we have proven ∎

Example 10.124. Let's consider the case that $G = B\mathcal{D}_3$ and verify the class equation in this special case. There are 6 conjugacy classes in this group (as usual, $\chi = \chi_6$):

$$\{1\} \qquad \{-1\} \qquad \{\chi, \chi^{-1}\}$$
$$\{\chi^2, \chi^{-2}\} \qquad \{\mathbf{j}, \chi^2\,\mathbf{j}, \chi^{-2}\,\mathbf{j}\} \qquad \{\chi\,\mathbf{j}, -\mathbf{j}, \chi^{-1}\,\mathbf{j}\}.$$

Therefore we can take $\mathbf{K}$ as $\{\chi, \chi^2, \mathbf{j}, \chi\,\mathbf{j}\}$. The centralizer of χ and χ^2 have 6 elements, while the centralizer of $\mathbf{j}$ and $\chi\mathbf{j}$ both only have 4 elements in their centralizers (recall that centralizers are cyclic in quaternion groups). Therefore we have

$$\begin{aligned} |\mathcal{C}(B\mathcal{D}_3)| + \frac{|B\mathcal{D}_3|}{|C(\chi)|} + \frac{|B\mathcal{D}_3|}{|C(\chi^2)|} + \frac{|B\mathcal{D}_3|}{|C(\mathbf{j})|} + \frac{|B\mathcal{D}_3|}{|C(\chi\,\mathbf{j})|} &= 2 + \frac{12}{6} + \frac{12}{6} + \frac{12}{4} + \frac{12}{4} \\ &= 2 + 2 + 2 + 3 + 3 \\ &= 12. \end{aligned}$$

Therefore the class equation checks out. A nice thing about quaternion groups, in this case $B\mathcal{D}_3$, is that you can see these 4 conjugacy classes as axis of rotation, see Figure 10.3. (This picture shows that there are always a total of 4 non-central conjugacy classes, one for an axis (from the origin) to the a vertex (n elements in cc), one from the origin to an edge (also n elements in cc), and one from origin (along the y-axis) pointing up and the last one pointing down along that same y-axis (these last two cc's contain $n-1$ elements each).

10.9.6.2 The Orbit Stabilizer Theorem for p-groups, OPST

Theorem 10.125 (The Orbit Stabilizer p–Theorem (**OPST**))**.** *Suppose that a p-group P acts on a finite set X. Then*

$$\#X \equiv_p \#X^P.$$

Proof. If a point $y \in X$ is not in the global fixed point set, X^P, of the action $P \circlearrowright X$, then there must be more than one point in the orbit of y, i.e. $\#P \bullet y > 1$. But by **OST**, we know that $\#P \bullet y$ is a divisor of $|P|$. Since P is a p–group, these last two facts mean that $\#P \bullet y = p^r$ for some integer r which is greater than 0. In particular, p divides the size of the orbit of any point not in the global fixed point set of the action, and so $0 \equiv_p \#(P \bullet y)$ for such points. Using this, and going back to (27) on 581, shows us that mod p, we have that

$$\#X = \#X^P + \sum \cancel{\#(P \bullet y)} \equiv_p \#X^P$$

as desired. ∎

Remark 10.126. An informal reason why the above works is that

This innocent looking lemma is one of the most useful facts in all of group theory! Let's give a more streamlined proof of Theorem 10.123.

Theorem 10.127 (=Theorem 10.123)**.** *If P is a p-group, then P is not centerless.*

Proof. (**Another proof of Theorem 10.123**) The global fixed point set of the action of P on itself by conjugation is the center of P, $\mathcal{C}(P)$. Thus

$$\cancel{|P|}^{0} \equiv_p |\mathcal{C}(P)|.$$

Hence p must divide $|\mathcal{C}(P)|$. Since, as for any group, $|\mathcal{C}(P)| \geq 1$, and 1 is not divisible by p, we must have that $|\mathcal{C}(P)| > 1$, i.e. $\mathcal{C}(P)$ is non-trivial. ∎

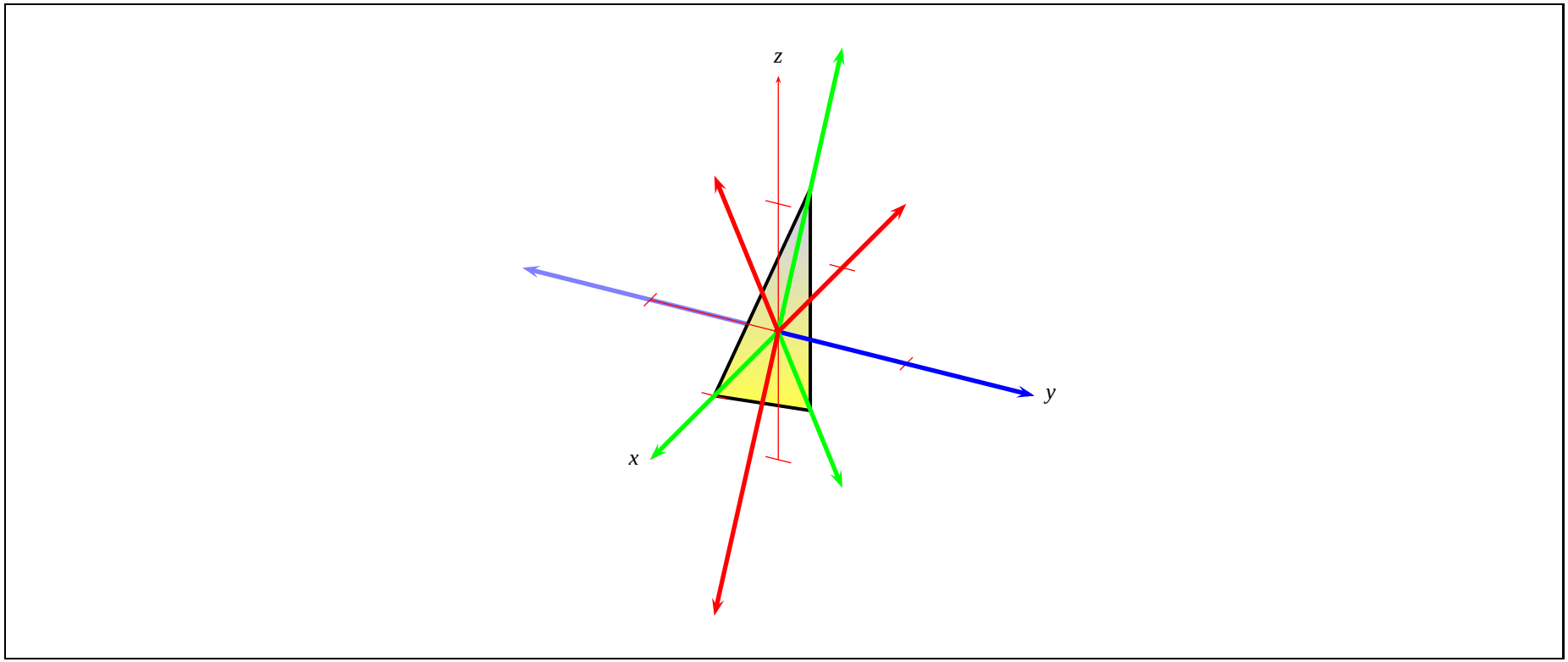


Figure 10.3: The 4 conjugacy classes of non-central elements in $B\mathcal{D}_3$

EXERCISE 10.68. Suppose that P is a p–group and suppose that N is a non-trivial normal subgroup of P Prove that $N \cap \mathcal{C}(P)$ is non-trivial.

Hint! Consider the action $P \circlearrowleft N$ by conjugation, $g \bullet n \stackrel{\text{def}}{=} gng^{-1}$. (Why is this an action?) Prove the global fixed point set of this action is $\mathcal{C}(P) \cap N$ to conclude via **OPST**

SOLUTION

$$|N| \equiv_p |\mathcal{C}(P) \cap N|$$

a proceed as above.

This lemma even as applications when the groups involved are not p-groups. In fact the proofs of all the Sylow's theorems will hinge on this little lemma. Here is an example of an incredibly powerful theorem which uses **OPST**.

REMARK 10.128.

(1) In the proof of the following recall that for a group G, and a positive integer n, $\mathbf{o}_n(G)$ is simply the number of elements of order n in G and $\mathbf{O}_n(G)$ is the number of elements whose order divides n.

(2) The original proof was given by Cauchy in 1845. However, we give a different proof. The proof we are about to present is one of the first real 'right-brained' proofs we've given so far! This specific proof, published 114 years later in 1959, was first given in [**?**] by James McKay. It is quite ingenious and its trajectory is certainly not something that should feel natural to you! In fact, the proof almost feels like a magic trick in many ways! It is one of the few places, in a first year group theory course, where credit is given not only to the originator of the theorem, Cauchy, but to the one who upgraded Cauchy's proof over 100 years later.

THEOREM 10.129 (Cauchy's Lemma[XV]). *Let p be a prime number which divides $|G|$. Then G contains an element of order p.*

PROOF. Let

$$\mathbf{X} = \{(g_1, \dots, g_p) : g_i \in G, g_1 \cdots g_p = e_G\} \subseteq G^p.$$

Now the cyclic group of order p acts on this set by resticting the action of the cyclic group of order p (written multiplicatively as $\langle \chi \rangle = \{\{p\}\}$ where χ is the standard element of order p to make things look more suggestive.) from Example { find }, i.e. by cyclically permuting the entries of G^p. Then we have

$$\chi^i \bullet (g_1, \dots, g_p) = (g_{1+i}, \dots, g_{p+i})$$

is still in $\mathbf{X}$ (subscripts are to be taken mod p). This is because

$$\begin{aligned}
(g_1, g_2, \dots, g_p) \in \mathbf{X} &\Rightarrow g_1 g_2 \cdots g_p = e_G && \text{(Def'n of } \mathbf{X}\text{)} \\
&\Rightarrow g_1^{-1} g_1 g_2 \cdots g_p = g_1^{-1} && \text{(Mul bs on left by } g_1^{-1}\text{)} \\
&\Rightarrow g_2 \cdots g_p = g_1^{-1} && \text{(Cancel } g^{-1}g\text{)} \\
&\Rightarrow g_2 \cdots g_p g_1 = g_1^{-1} g_1 && \text{(Mul bs on right by } g_1\text{)} \\
&\Rightarrow g_2 \cdots g_p g_1 = e_G && \text{(Cancel } g^{-1}g\text{)} \\
&\Rightarrow (g_2, \dots, g_p, g_1) \in \mathbf{X} && \text{(Def'n of } \mathbf{X}\text{)}.
\end{aligned}$$

But this last quantity, $(g_2, \dots, g_p, g_1)$ is by definition $\chi \bullet (g_1, g_2, \dots, g_p)$. Therefore, for an arbitrary $\mathbf{g} \in \mathbf{X}$, we have proven that $\chi \bullet \mathbf{g}$ is still in $\mathbf{X}$. Now (inductively) we also have for all integers i, $\chi^i \bullet \mathbf{g}$ is still in $\mathbf{X}$. In other words, the action of $\{\{p\}\} \circlearrowright G^p$ preserves $\mathbf{X}$, and so we also have an action $\{\{p\}\} \circlearrowright \mathbf{X}$.

Since $\{\{p\}\}$ is a p-group (it actually has order exactly equal to p), **OPST** implies that

$$\#\mathbf{X}^{\{\{p\}\}} \equiv_p \#\mathbf{X}. \tag{29}$$

We evaluate both sides of this equation by showing

Step 1 $\#\mathbf{X}^{\{\{p\}\}} = \mathbf{o}_p(G) + 1$, and

Step 2 p divides $\mathbf{X}$.

Let's get started. (Woooot!)

Step 1 Now if $(g_1, \dots, g_p) \in \mathbf{X}^{\{\{p\}\}}$, then by the definition of this, we have

$$(g_1, g_2, \dots, g_p) = \chi \bullet (g_1, g_2, \dots, g_p) = (g_2, \dots, g_p, g_1).$$

[XV] Often this is called Cauchy's *Theorem* instead of Lemma.

Hence, by definition of a tuple, each of the coordinates must be equal. But the first coordinate of the first tuple is g_1 and the first coordinate of the second tuple of g_2. Hence $g_1 = g_2$. At the same time the second component of the first tuple of g_2 while the second component of the second tuple is g_3. Thus $g_2 = g_3$. But since we've already shown that $g_1 = g_2$, we have that $g_1 = g_3$. Continuing in this way we see that $g_1 = g_i$ for each integer $i = 1 \dots p$. That is, if we let g equal g_1 we have that

$$(g, \dots, g) = (g_1, \dots, g_p).$$

We still haven't used the fact that $(g_1, \dots, g_p) \in \mathbf{X}$ in this part. Well let's do this

$$\begin{aligned} (g_1, \dots, g_p) \in \mathbf{X} &\Rightarrow g_1 \cdots g_p = e_G && \text{(Def'n of } \mathbf{X}\text{)} \\ &\Rightarrow g \cdots g = e_G && \text{(Since each } g_i = g\text{).} \\ &\Rightarrow g^p = e_G. \end{aligned}$$

In conclusion,

$$\mathbf{X}^{\{[p]\}} = \{(g, \dots, g) : g^p = e_G\}.$$

This (corresponds with) is the set of elements whose p^{th} power is the identity. Since p is assumed prime, the only such elements are the identity and elements of order p. In conclusion,

$$\#\mathbf{X}^{\pi_p} = \mathbf{o}_p(G) + 1. \tag{30}$$

Step 2 Next we show that $\#\mathbf{X} = |G|^{p-1}$ by constructing a bijection, in particular, since we are assuming that p divides $|G|$, this implies that $p \mid \#\mathbf{X}$. For $(g_1, \dots, g_p) \in \mathbf{X}$, we have, by definition of $\mathbf{X}$, $g_1 \cdots g_p = e_G$. As such, we must have $g_p = \left(g_1 \cdots g_{p-1}\right)^{-1}$. Conversely, given any element $(g_1, \dots, g_{p-1})$ of G^{p-1}, we have that

$$\left(g_1, g_2, \dots, g_{p-1}, (g_1 \cdots g_{p_1})^{-1}\right) \in \mathbf{X}.$$

Therefore, we have proven that

$$\begin{aligned} &: G^{p-1} \to \mathbf{X} \\ &: (g_1, \dots, g_{p-1}) \mapsto \left(g_1, g_2, \dots, g_{p-1}, (g_1 \cdots g_{p_1})^{-1}\right) \end{aligned}$$

is a bijection. Therefore, we have shown what we wanted, namely that $\#\mathbf{X} = |G|^{p-1}$ and as pointed out this implies that

$$p \mid \#\mathbf{X}. \tag{31}$$

Now using (30) and (31) to plug in for both sides of (29) we see

$$0 \equiv_p \mathbf{o}_p(G) + 1 \Rightarrow -1 \equiv_p \mathbf{o}_p(G). \tag{32}$$

In particular, since -1 does not equal 0 mod p, we have that $\mathbf{o}_p(G) \neq 0$, that is there exists an element of order p in G, as desired. ∎

For instance, this says that the group $\mathbf{SL}_3(\mathbb{Z}_3)$ which we showed has order $5,616 = 2^4 \cdot 3^3 \cdot 13$ must have an element of order 13. Of course finding this element is a different matter. [XVI]

[XVI] Actually the matrix

$$\begin{pmatrix} [0]_3 & [0]_3 & [1]_3 \\ [1]_3 & [0]_3 & [0]_3 \\ [0]_3 & [1]_3 & [2]_3 \end{pmatrix}$$

works (i.e. has order 13). I was able to find this matrix right away by using some stuff from algebraic number theory which one sees in the second year as a graduate student. If you have nothing better to do for the next few ... I guess 20 minute intervals ... feel free to verify this has order 13.

WARNING! 10.130. Cauchy's Lemma does not work when p is not a prime! Indeed, only cyclic groups have an element of order whose order matches any given divisor of the order of the group.

For instance the smallest non-cyclic group, which is $\mathbb{Z}_2 \times \mathbb{Z}_2$, does not have an element of order 4.

In fact for many groups, there are many divisors of the group's order with no corresponding elements of that order. For instance, every non-identity element of $\mathcal{A}_4$ has order 2 or 3 and none have order $2^2 = 4$ (or 6).

For a more extreme case, we simply need to look at $\mathcal{A}_n$ for higher n. For instance, the order of $\mathcal{A}_5$ is 60, which is a number with 12 divisors. However $\mathcal{A}_5$ has only elements of order 1, 2, 3, or 5. The order of $\mathcal{A}_6$ (which is 360) has 24 divisors, but only has elements of the orders listed for $\mathcal{A}_5$ and 4.

We have actually proven something a bit stronger than what we said.

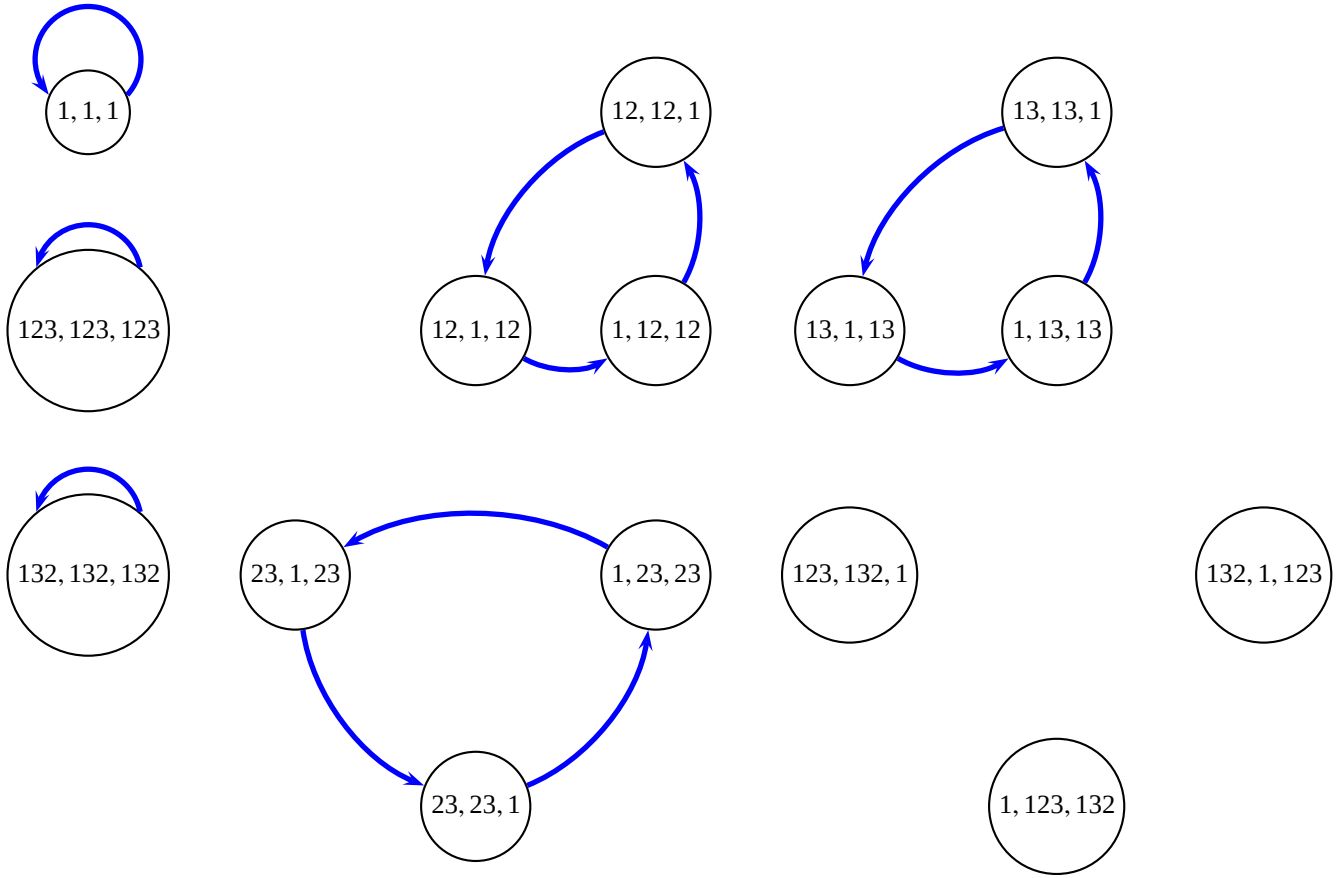
COROLLARY 10.131. *Let p be a prime dividing the order of a finite group G. Then $\mathbf{o}_p(G) \equiv_p -1$.*

PROOF. Check out the proof and note that this is actually what we proved there. Indeed, Equation (32) says exactly this. ∎

The following Corollary puts together Lagrange's theorem and Cauchy's lemma. It is a tremendous generalization of Exercise 5.35 back on page 270.

COROLLARY 10.132. *A finite group has an element of order p (p prime) iff $p \mid |G|$.* ∎

EXAMPLE 10.133. Let's work an example which follows the proof of Cauchy's lemma. Let's take $G = \mathcal{S}_3$ and $p = 3$. Then there are $6^{3-1} = 6^2 = 36$ elements in $\mathbf{X}$ (here $\mathbf{X}$ is as defined in the proof of Cauchy's lemma). Let's draw the action diagram {Still need to connect the dots! }



EXERCISE 10.69. Prove that there is a 3 by 3 matrix with determinant 1 and coefficients in $\mathbb{Z}_5$ whose order is 31.

SOLUTION

EXERCISE 10.70. Prove that there are a total of 806 subgroups of $\mathbb{Z}_5^4$ which are isomorphic to $\mathbb{Z}_5^2$.[XVII]

Hint! Show that $\mathbf{GL}_4(\mathbb{Z}_5)$, which we think of as $\mathrm{Aut}\left(\mathbb{Z}_5^4\right)$, acts transitively on the set of such subgroups.[XVIII](For generators x and y of a subgroup of $\mathbb{Z}_5^4$ isomorphic to $\mathbb{Z}_5^2$ take a matrix whose first column is x and whose second column is y to move $\begin{bmatrix}[1]_5\\[0]_5\\[0]_5\end{bmatrix} \mapsto x$ and $\begin{bmatrix}[0]_5\\[1]_5\\[0]_5\end{bmatrix} \mapsto y$)

SOLUTION

Next show that the stabilizer of

$$\left\langle \begin{bmatrix} [1]_5 \\ [0]_5 \\ [0]_5 \end{bmatrix}, \begin{bmatrix} [0]_5 \\ [1]_5 \\ [0]_5 \end{bmatrix} \right\rangle$$

in $\mathbf{GL}_4(\mathbb{Z}_5)$ is given by the set of invertible matrices of the form

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ [0]_5 & [0]_5 & A_{33} & A_{34} \\ [0]_5 & [0]_5 & A_{43} & A_{44} \end{pmatrix}.$$

Show that the order of this stabilizer is $5^4 \cdot |\mathbf{GL}_2(\mathbb{Z}_5)|^2$. Then, using **OST**, conclude there are

$$\frac{|\mathbf{GL}_4(\mathbb{Z}_5)|}{5^4 \cdot |\mathbf{GL}_2(\mathbb{Z}_5)|^2}$$

such subgroups.

Exercise 10.71. Generalize the previous exercise to show that there are a total of

$$\frac{(p^n - 1)(p^{n-1} - 1)}{(p^2 - 1)(p - 1)}$$

subgroups of $\mathbb{Z}_p^n$ of order $\mathbb{Z}_p^2$.

Solution

Exercise 10.72. Inspect the proof of Cauchy's lemma. What step fails if instead we try to follow the same proof to show that if $p^2 \Big| |G|$, then there exists an element of order p^2 in G? (There is only one subtle step that fails!)

Solution

Exercise 10.73. Let G be any finite group. Prove for any prime p, p divides the number of elements of order p^2 in G.

Solution

[xvii] Any two non-identity elements of the bigger group, $\mathbb{Z}_5^4$, which are not multiples of each other do the job.

[xviii] Of course you can write the action on subgroups as

$$\mathbf{A} \bullet \langle x, y \rangle := \langle \mathbf{A}(x), \mathbf{A}(y) \rangle .$$

Exercise 10.74. Let p and q be primes and $n > 1$ be an integer.

(1) Prove there exists a matrix of degree n and coefficients in $\mathbb{Z}_p$ of order q iff there exists a matrix of determinant 1 with coefficients in $\mathbb{Z}_p$ of order q.

(2) (Now we show that this doesn't work for integers not a prime.) Show that $\mathbf{GL}_2(\mathbb{Z}_3)$ contains an element of order 8, but $\mathbf{SL}_2(\mathbb{Z}_3)$ does not.
Hint! Perhaps the fastest way to see that $\mathbf{SL}_2(\mathbb{Z}_3)$ does not have an element of order 8 is to recall that this group is isomorphic to $B\left(\blacktriangleright^3\right)$, and this group does not have an element of order 8 since $S\left(\blacktriangleright^3\right) \cong \mathcal{A}_4$ does not have an element of order 4.

SOLUTION

10.9.7 Groups of twice a prime

Theorem 10.134. *Let p be a prime. Then every group of order $2p$ is either $\mathcal{D}_p$ or $\mathbb{Z}_{2p}$.*[XIX]

Proof. Let G be a group of this order. We have already classified groups of order 4, so we assume that G does not have order 4. That is we are assuming that p is an odd prime. Then by Cauchy's lemma, we know that G contains elements x of order p and y of order 2. Since $|x|$ and $|y|$ are relatively prime, we know that $G = \langle x, y \rangle$. Since $\langle x \rangle$ has order $p = |G|/2$, we can apply Theorem 10.61 to conclude that $\langle x \rangle$ is a normal subgroup of G. In particular, we know that

$$yxy^{-1} \in \langle x \rangle \Rightarrow yxy^{-1} = x^i$$

for some integer $0 \leq i < p$. Moreover, since y has order 2 we also have that

$$x = y^2xy^{-2} = yyxy^{-1}y^{-1} = yx^iy^{-1} = \left(yxy^{-1}\right)^i = \left(x^i\right)^i = x^{i^2} \Rightarrow i^2 = 1 \mod p.$$

Since p is prime number, division makes sense in $\mathbb{Z}_p$ for all non-zero elements and as such there are no zero divisors [XX], so we know that

$$i^2 - 1 = (i-1)(i+1) = 0 \Rightarrow i - 1 = 0 \text{ or } i + 1 = 0 \Rightarrow i = 1 \text{ or } i = -1.$$

So let's deal with these two cases individually:

$i = +1$ Here, $yxy^{-1} = x^1 = x$ and so x and y commute. Moreover since their orders are relatively prime, we also have

$$\langle x \rangle \cap \langle y \rangle = \{e_G\}.$$

Therefore, we can apply Lemma 5.58 to conclude that $|xy| = 2p$. As this is the order of G we know that G is therefore isomorphic to $\mathbb{Z}_{2p}$.

$i = -1$ Now we can apply Theorem 10.81 to conclude that G is isomorphic to $\mathcal{D}_{2p}$ and the proof is complete. ■

At this time, you should run, not walk, to Appendix I which starts on page 865 to verify this theorem! For instance, this table shows us that there are 2 groups of orders 4, 6, 10, 14, 22, 26 ... etc, etc.

[XIX] Again, we consider a rectangle which is not a square a regular 2-gon. This BS consideration just makes this theorem easier to state!
[XX] Recall that a ring has zero divisors if there exist non-zero elements r and s whose product is zero

10.9.7.1 Applications to $\mathcal{A}_5$

Lemma 10.135. *$\mathcal{A}_5$ contains no proper subgroup of order greater than 12.*

Proof. The group $\mathcal{A}_5$ is a simple group. We prove this in a bit, Theorem 14.11. Below I give an outline of how to prove this without using this and leave it for an exercise for you to fill in the details. You can take this fact for granted for now. Now suppose that there were a subgroup H of $\mathcal{A}_5$ of order greater than 12. Then its index in $\mathcal{A}_5$ would be less than $60/12 = 5$. As such, we'd have a non-trivial homomorphism $\rho\colon \mathcal{A}_5 \to \mathcal{S}_{\mathcal{A}_5/H}$, which is a representation of the coset action,

$$\mathcal{A}_5 \circlearrowright \mathcal{A}_5/H$$
$$\sigma \bullet \tau H := \sigma\tau H.$$

Since H has index less than 5, $\mathcal{S}_{\mathcal{A}_5/H}$ has 4! or fewer elements. Thus, since the order of $\mathcal{A}_5$ is bigger than 4!, ρ has to have a proper (since the action is transitive, hence its kernel cannot be everything) non-trivial (since the homomorphism is not injective) kernel. Since the kernel of any homomorphism is a normal subgroup, we've proven that $\mathcal{A}_5$ has a PNT, normal subgroup. But this is false since $\mathcal{A}_5$ is simple. ■

Solution

Exercise 10.75. Follow the following steps to prove that the above lemma without using the fact that $\mathcal{A}_5$ is simple.

(1) Prove that

$$\mathcal{A}_5 = \langle (1\,2\,3\,4\,5), (2\,1\,3\,4\,5) \rangle$$

Hint! We know that $(1\,2)$ and $(1\,2\,3\,4\,5)$ generate $\mathcal{S}_5$. Use this and the fact that

$$(1\,2)(1\,2\,3\,4\,5)(1\,2)^{-1} = (2\,1\,3\,4\,5),$$

the second of the two listed generators.
Careful! Of course, Theorem 10.136 below implies this, but since this exercise is a lemma to prove that theorem you certainly cannot use it to prove this exercise!

(2) Prove that the image of any non-trivial homomorphism from $\mathcal{A}_5$ must be generated by elements of order 5.

(3) Prove that a non-trivial homomorphism $\mathcal{A}_5 \to \mathcal{S}_4$ does not exist.[XXI]
Hint! $\mathcal{S}_4$ does not contain an element of order 5.

(4) Prove that a subgroup of index less than 5 of $\mathcal{A}_5$ gives rise to a non-trivial homomorphism $\mathcal{A}_5 \to \mathcal{S}_4$.
Hint! This much was explained inside of the proof of the lemma.

(5) Conclude (like we did in the lemma) that no such subgroup can exist.

Theorem 10.136. *Suppose σ and τ are $\underline{5}$–cycles in $\mathcal{A}_5$ which are not powers of each other. Then together σ and τ and generate all of $\mathcal{A}_5$.*

Proof. The assumption can be reworded to say that

$$\langle \sigma \rangle \cap \langle \tau \rangle = \{(1)\}.$$

Hence we have that

$$\langle \sigma \rangle \cap \langle \tau \rangle = \{\sigma^i \tau^j : i, j \in \mathbb{Z}\}$$

[XXI] Again, the point of this exercise is that you are not allowed to use the fact that $\mathcal{A}_5$ is simple.

has order $5 \cdot 5 = 25$. But any subgroup of $\mathcal{A}_5$ which contains σ and τ must contain $\langle \sigma \rangle \cap \langle \tau \rangle$. Hence if σ and τ did not generate all of $\mathcal{A}_5$ we'd have a proper subgroup of $\mathcal{A}_5$ of order greater than 25. This is impossible by the lemma. ■

EXERCISE 10.76. Let σ be a 5–cycle and τ be a 3–cycle in $\mathcal{A}_5$. Prove that σ and τ generate all of $\mathcal{A}_5$.
Hint! Follow the above proof but instead prove that the group generated by these two permutations has order at least 15.

SOLUTION

EXERCISE 10.77. Show that any 5–cycles and any 2–cycle generates $\mathcal{S}_5$.
Hint! If you denote the 2–cycles as τ and the 5–cycles as σ we have that $\tau\sigma\tau^{-1}$ is a 5–cycle. Show that this 5–cycles is not a power of σ. Therefore $\left\langle \sigma, \tau\sigma\tau^{-1} \right\rangle = \mathcal{A}_5$ by Theorem 10.136. But since τ and both σ and $\tau\sigma\tau^{-1}$ are even, we have that

$$\tau \notin \left\langle \sigma, \tau\sigma\tau^{-1} \right\rangle.$$

Therefore, $\langle \sigma, \tau \rangle$ must be a subgroup of $\mathcal{S}_5$ properly containing $\mathcal{A}_5$.

SOLUTION

10.10 GENERATORS OF STABILIZERS*

We have seen that it seems at least reasonably difficult to compute a generating set for a subgroup of a group, even if we are given a perfectly group generating set for the larger group.

EXAMPLE 10.137. To motivate things without driving you nuts with notation, I begin by presenting a very very special case: suppose G can be generated by two elements, $G = \langle x, y \rangle$, and H is an index 2 subgroup of G which contains x but does not contain y. Since H is index 2, we know that despite the fact that $y \notin H$, we know that $y^2 \in H$. Moreover since any index 2 subgroup is normal, we also know that $yxy^{-1} \in H$. In turns out, together with x itself, there are the only 3 elements we need to throw in to get a generating set for H.

EXAMPLE /THEOREM 10.138. In the case here, H is generated by y^2, x, and yxy^{-1}.

PROOF. Let

$$\widehat{H} := \left\langle y^2,\ x,\ y^{-1}xy \right\rangle.$$

Of course $\widehat{H} \subseteq H$, we wish to show the reverse equality also holds. We do this by showing that our new friend, $\widehat{H}$ also has index 2 in G. Indeed, we know that since $y \notin H$, we know that $y \notin \widehat{H}$ as well. Therefore $\widehat{H}$ and $y\widehat{H}$ are distinct left cosets of $\widehat{H}$ in G. Let's denote $\mathbf{C} := \left\{\widehat{H}, y\widehat{H}\right\}$. We show that $\mathbf{C} = G/\widehat{H}$ (proving that the index of $\widehat{H}$ in G is 2) by proving that the transitive action $G \circlearrowright G/\widehat{H}$ leaves $\mathbf{C}$ invariant and then using Exercise 10.13.

Reducing the problem further, since the set of elements of G which leave $\mathbf{C}$ invariant, $G_{\mathbf{C}}$, is a subgroup of G, and G is generated by x and y, we simply need to check that $x, y \in G_{\mathbf{C}}$. Let's get started with this simple task.

(1) First let's check out what x does to the two cosets. Since x is one of the three listed generators of $\widehat{H}$, we know that $x \in \widehat{H}$ and as such $x \bullet \widehat{H} = \widehat{H}$. What does x do to

$y\widehat{H}$? (It better fix it as well!) Well, since $y^{-1}xy$ is one of the generators listed for $\widehat{H}$, we know that

$$\begin{aligned} y^{-1}xy \bullet \widehat{H} &= \widehat{H} \\ \Rightarrow xy \bullet \widehat{H} &= y\widehat{H} \\ \Rightarrow x \bullet (y\widehat{H}) &= y\widehat{H} \end{aligned}$$

and so x also fixes the second coset $y\widehat{H}$ and leaves $\mathbf{C}$ invariant (in a very strong way!).

(2) Next, we need to check to see what y does to $\mathbf{C}$. Well, by definition, we know that $y \bullet \widehat{H} = y\widehat{H}$. Therefore, we simply need to check what y does to $y\widehat{H}$. (It better send it back to $\widehat{H}$!) Well, since y^2 is one of the three generators listed for $\widehat{H}$, we know that $y^2\widehat{H} = \widehat{H}$. But this means that $y \bullet (y\widehat{H}) = y\widehat{H}$ as desired. ■

Let's use this in a specific case. Let's take $G = \mathcal{S}_7$ and $H = \mathcal{A}_7$. We know that $\mathcal{S}_7$ is generated by $(1\,2)$ and $(1\,2\,3\,4\,5\,6\,7)$. The former permutation is odd (and hence not in H) and the latter is even (and hence is in H). Therefore, by the example/theorem we know that

$$\mathcal{A}_7 = \left\langle \cancel{(1\,2)^2}, (1\,2\,3\,4\,5\,6\,7), (1\,2)(1\,2\,3\,4\,5\,6\,7)(1\,2)^{-1} \right\rangle = \langle (1\,2\,3\,4\,5\,6\,7), (2\,1\,3\,4\,5\,6\,7) \rangle.$$

Our special case is a bit more useful than it may first appear. For instance, if we had an index 2 subgroup H with neither x nor y in H, then since H has index 2, we know for a fact that $yx \in H$. Therefore, we can repeat the above (with yx replacing x) to be able to prove that H can be generated by $\{y^2, yx, xy\}$. For instance, we know that $\mathcal{S}_8$ can be generated by $(1\,2\,3\,4\,5\,6\,7\,8)$ and $(1\,2)$. Since both permutations are odd, neither is in $\mathcal{A}_8$. However using our new tool, we know that

$$\mathcal{A}_8 = \langle \cancel{(1\,2)^2}, (1\,2)(1\,2\,3\,4\,5\,6\,7\,8), (1\,2\,3\,4\,5\,6\,7\,8)(1\,2) \rangle = \langle (2\,3\,4\,5\,6\,7\,8), (1\,3\,4\,5\,6\,7\,8) \rangle.$$

EXERCISE 10.78.

SOLUTION

(1) If n is odd, prove that $\mathcal{A}_n$ can be generated by two n–cycles.

(2) If n is even, prove that $\mathcal{A}_n$ can be generated by two $n-1$–cycles.

We considered the special case above not as a lemma of any sort, but as a motivator. We will see that, rather surprisingly, the proof of the more general theorem will go by faster than the proof of the special case. That's why it's a good idea to get to know the special case a bit better since the notation in the proof of the general theorem may become overwhelming without a concrete example to keep us grounded.

Notice that our Example/Theorem we were, in a sense, led to what we needed to throw into a generating set for H by noting that at each step in order for a generating to leave $\mathbf{C}$ invariant, we were required to have a certain element of H in there. As you saw in the Example/Theorem, it's easiest to consider the general theorem in terms of actions. Since every subgroup of a group is the stabilizer of some action, namely the coset action, we are not losing anything by considering transitive actions. We now move onto the general case.

Suppose $G \circlearrowright X$ is a transitive action of a group G which can be generated by a set S on a finite set X, $\{x_1, \ldots, x_n\}$. Let's denote x_1 as x. For any $i \in \overline{n}$, since the action $G \circlearrowright X$ is transitive there exists at least one element of G which takes x to x_i. Let's pick one such element of G and denote it t_i. Furthermore, let's denote the collection of these t_i as $\mathcal{T}$. That is we are assuming that

$$t_i \bullet x = x_i$$

and defining $\mathcal{T} := \{t_i : i \in \overline{n}\}$. Moreover, since every $g \in G$ must do something to x, we have that

$$g \bullet x = t \bullet x_m$$

for a unique $t \in \mathcal{T}$. Let's denote this t as T_g. Lastly, for a pair $(t, s) \in \mathcal{T} \times S$ let's define

NOTATION 10.139.

$$\tau(s,t) \stackrel{\text{def}}{=} (st)^{-1}T_{st}.$$

Note that by definition, we have for any $g \in G$

$$g \bullet x = T_g \bullet x \Rightarrow g^{-1}T_g \in G_x.$$

In particular, since each $\tau(t,s)$ is of this form (with $g = st$) we have that $\tau(s,t) \in G_x$. The content of the next theorem is that these dudes actually form a generating set for G_x.

THEOREM 10.140. *Let G, S, $\mathcal{T}$, and X be as above. Then G_x can be generated by*

$$\{\tau(s,t) \mid t \in \mathcal{T}, s \in S\}. \tag{33}$$

PROOF. Denote G_x as H and $\widehat{H}$ as the group generated by (33). We wish to prove that $\widehat{H} = H$. In the paragraph before the proof, we have already argued that each $\tau(s,t) \in H$. So we have one inclusion, $\widehat{H} \subseteq H$, right off the bat. Let's now prove that $H \subseteq \widehat{H}$.

Note that $\mathcal{T}$ is nothing more than a set of coset representatives for G/H. Copying the proof of the special case from Example/Theorem 10.138 we simply need to prove that the collection of left cosets

$$\mathbf{C} := \left\{t\widehat{H} : t \in \mathcal{T}\right\}$$

is invariant under the action of each element of the generating set S.

To this end, let $s \in S$ and $t \in \mathcal{T}$ be arbitrary. We need to prove that $s \bullet t\widehat{H} \in \mathbf{C}$. Since it's one of the listed generators,

$$\tau(s,t) = (st)^{-1}T_{st} \in \widehat{H}.$$

Therefore

$$st\tau(s,t) = T_{st} \Rightarrow st\widehat{H} = T_{st}\widehat{H}.$$

But $T_{st} \in \mathcal{T}$ and so $T_{st}\widehat{H} \in \mathbf{C}$ completing the proof. ∎

Typically, we set up a table to help guide us. We first pick our set $\mathcal{T}$ and label the rows with the elements of this set. We label the columns with our generating set S. Of course the answer we get will depend on our choices, so don't be afraid to try several combinations. For example, we can always assume that t_1 is simply the identity e. Then we fill in our table by filling in the entry located at the i^{th} row and j^{th} column with $\tau(s_j, t_i)$. After completing the table, we entries will be our generating set.

EXAMPLE 10.141. Let G be the group

$$\left\langle (1\,2\,3\overset{x}{\overset{\shortparallel}{4}}5\,6\,7), (1\,2)\overset{y}{\overset{\shortparallel}{(}}3\,6) \right\rangle$$

considered in Example 2 back on page 502. Let's find a generating set for G_1, the stabilizer of the letter $1 \in \overline{7}$. Since there is a unique power of x which takes 1 to any letter in $\overline{7}$, we take $\mathcal{T}$ as the set $\langle x \rangle$.[1]

[1] There is no group structure on $\mathcal{T}$ in general. The notation used is simply used for convenience.

Let's set up a table to help us:

x^6	
x^5	
x^4	
x^3	
x^2	
x	
(1)	
	x y

.

Notice that we have $tt^i = t^{i+1} \in \mathcal{T}$. Therefore, $\tau(t, t^i) = (1)$ and we can fill in our entire first column with just (1)'s:

x^6	(1)
x^5	(1)
x^4	(1)
x^3	(1)
x^2	(1)
x	(1)
(1)	(1)
	x y

.

As you should expect, the second column is more interesting. Let's get started:

$$
\begin{aligned}
y \bullet 1 = 2 = x \bullet 1 \quad &\Rightarrow (y)^{-1} \circ x = yx = \tau(y,(1)) = x^0 y x^1 \\
yx \bullet 1 = y \bullet 2 = 1 = (1) \bullet 1 \quad &\Rightarrow (yx)^{-1} \circ (1) = x^{-1}y = \tau(y,x) = x^6 y x^0 \\
yx^2 \bullet 1 = y \bullet 3 = 6 = x^5 \bullet 1 \quad &\Rightarrow (yx^2)^{-1} \circ x^5 = x^{-2} y x^5 = \tau(y,x^2) = x^5 y x^5 \\
yx^3 \bullet 1 = y \bullet 4 = 4 = x^3 \bullet 1 \quad &\Rightarrow (yx^3)^{-1} \circ x^3 = x^{-3} y x^3 = \tau(y,x^3) = x^4 y x^{-4} \\
yx^4 \bullet 1 = y \bullet 5 = 5 = x^4 \bullet 1 \quad &\Rightarrow (yx^4)^{-1} \circ x^4 = x^{-4} y x^4 = \tau(y,x^4) = x^3 y x^{-3} \\
yx^5 \bullet 1 = y \bullet 6 = 3 = x^2 \bullet 1 \quad &\Rightarrow (yx^5)^{-1} \circ x^2 = x^{-5} y x^2 = \tau(y,x^5) = x^2 y x^2 \\
yx^6 \bullet 1 = y \bullet 7 = 7 = x^6 \bullet 1 \quad &\Rightarrow (yx^6)^{-1} \circ x^6 = x^{-6} y x^6 = \tau(y,x^6) = xyx^{-1}
\end{aligned}
$$

So the table is completed as follows:

x^6	(1)	xyx^{-1}
x^5	(1)	x^2yx^2
x^4	(1)	x^3yx^{-3}
x^3	(1)	x^4yx^{-4}
x^2	(1)	x^5yx^5
x	(1)	x^6y
(1)	(1)	yx
	x	y

.

Let's label these 7 elements of G_1 as h_1 through h_7 starting with $h_1 := yx$ and ending with $h_7 := xyx^{-1}$. Let's denote the set of these 7 h's as $\mathfrak{H}$. The theorem says that $\langle \mathfrak{H} \rangle = G_1$. We want to throw away some elements of $\mathfrak{H}$ to reduce $\mathfrak{H}$ to a minimal generating set of G_1, i.e. a generating set which does not properly contain another generating set. Note that

$$h_2h_1 = (x^6y)(yx) = x^7 = (1) \Rightarrow h_2^{-1} = h_1.$$

Therefore we can get rid of one of them, say h_2 and redefine $\mathfrak{H}$ as $\mathfrak{H} \setminus \{h_2\}$.

At the same time, we also have that

$$h_3h_6 = x^5yx^5x^2yx^2 = x^5y^2x^2 = x^5x^2 = (1) \Rightarrow h_6^{-1} = h_3$$

and we can also get rid of, say, h_6 and further redefine $\mathfrak{H}$ as $\mathfrak{H} \setminus \{h_6\}$. (So far we've only used the facts that $|x| = 7$ and $|y| = 2$).

Now let's explicitly compute the other five h_i's in $\mathfrak{H}$ and write them in cycle notation. The conjugacy ones are easiest since conjugation is replacement in $\mathcal{S}_n$. As such

$$h_7 = (2\,3)(4\,7) \qquad h_5 = x^2h_7x^{-2} = (4\,5)(6\,2) = (2\,6)(4\,5) \qquad h_4 = xh_3x^{-1} = (3\,7)(5\,6).$$

Next,

$$\begin{aligned} h_1 &= yx \\ &= (1\,2)(3\,6)(1\,2\,3\,4\,5\,6\,7) \\ &= (1)(2\,6\,7)(3\,4\,5) \\ &= (2\,6\,7)(3\,4\,5) \end{aligned}$$

$$\begin{aligned} h_3 &= x^3(x^2yx^{-2}) \\ &= (1\,4\,7\,3\,6\,2\,5)(3\,4)(5\,1) \\ &= (4\,7\,3\,6\,2\,5)(1\,5)(1\,5)(3\,4) \\ &= (4\,7\,3\,6\,2\,5)(3\,4) \\ &= (2\,5\,4\,6)(3\,7). \end{aligned}$$

Meanwhile

$$\begin{aligned} h_5h_4 &= (2\,6)(4\,5)(3\,7)(5\,6) \\ &= (2\,6)(4\,5\,6)(3\,7) \\ &= (2\,6\,4\,5)(3\,7) \\ &= h_3^{-1}, \end{aligned}$$

and so we also do not need $h_3 \in \mathfrak{H}$ and so we replace $\mathfrak{H}$ with $\mathfrak{H} \setminus \{h_3\}$.

At this point, we have cut down on the generating set of G_1, the stabilizer of the letter 1, significantly enough allow for some direct calculations in it. In fact, with a bit of effort it is not overly difficult to see that $G_1 \cong \mathcal{S}_4$ were we view $\mathcal{S}_4$ as a transitive subgroup of $\mathcal{S}_6$ by having it permute the 6 faces of a cube. If this isn't the way you'd

like to proceed, you can use the theorem again to now find the stabilizer of 2 in G_1 (so we would end up computing $G_{(1,2)}$) and from there we would, could, and should see that this stabilizer consists of the four elements

$$\{(1),\ h_4,\ h_3h_4h_3^{-1},\ h_3^2h_4h_3^{-2}\}$$

(which is a group isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$.) Since the orbit of G_1 consists of all 6 letters $\{2,3,4,5,6,7\}$[II] this would give us yet another proof that G_1 consists of 24 elements (and so the bigger group G consists of 168 elements.)

Instead, to avoid any further calculations, we use a bit more trickery to show that G_1 has 24 elements. Let's define $\sim$ as the equivalence relation corresponding to the partition

$$\widehat{1} := \{2,4\} \qquad \widehat{2} := \{5,6\} \qquad \widehat{3} := \{3,7\}.$$

Note that G_1 preserves this equivalence relation and

$$h_1 \bullet = (\widehat{1}\,\widehat{2}\,\widehat{3}) \qquad h_4 \bullet = (\widehat{1})(\widehat{2})(\widehat{3}) \qquad h_5 \bullet = (\widehat{1}\,\widehat{2})(\widehat{3}) \qquad h_7 \bullet = (\widehat{1}\,\widehat{3})(\widehat{2}).$$

Therefore, the image of $\varphi\colon G_1 \to \mathcal{S}_{\{\widehat{1},\widehat{2},\widehat{3}\}}$ is (isomorphic to) all of $\mathcal{S}_3$. For an element to be in the kernel of this homomorphism it must swap 2 and 4, 5 and 6, and 3 and 7. Since $G \leqslant \mathcal{A}_7$ (it is generated by even permutations) there are four possible ways of doing this in all of $G_1 \leqslant \mathcal{S}_{\{2,3,4,5,6,7\}} \leqslant \mathcal{A}_7$. Moreover all four occur, indeed,

$$h_4 = (3\,7)(5\,6) \qquad h_1h_4h_1^{-1} = (2\,4)(3\,7)$$

generate all four permutations which do just that. Therefore, the kernel of the map φ contains four elements and the image contains 6, therefore by **KIT** G_1 contains 24 elements.

We have finally proved that the action of G acts transitively on 7 letters (this much was obvious) and has a stabilizer containing 24 elements (this turned out to be not obvious!). Therefore, using **OST**, we have proven that

$$|G| = 7 \cdot 24 = 7 \cdot 3 \cdot 2^3 = 168.$$

[II] This reflects the fact that the action of $G \circlearrowright \overline{7}$ is 2–transitive.

11 Quotient Groups

So far have mainly been interested in malnormal subgroups of a group G. Now we will study normal subgroups of G. While malnormal subgroups give us a faithful way of turning a group into a symmetric group, normal subgroups will give us a way of simplifying a group. This gives us a powerful way of computing the order of a group as well as many other things!

11.1 Definitions

Recall, for any group G and for any subgroup $H \leqslant G$ we have an action $G \circlearrowright G/H$ given by $g \bullet aH = gaH$. This means that we have a function

$$\bullet \colon G \times {}^{G}\!/_{H} \to {}^{G}\!/_{H}.$$

This is pretty close to what we need to be able to define a binary operation from

$$\begin{aligned} \otimes \colon {}^{G}\!/_{H} \times {}^{G}\!/_{H} &\to {}^{G}\!/_{H} \\ gH \otimes aH &:= g \bullet aH = gaH. \end{aligned}$$

However, unlike the action, this is often **not** well defined, i.e. it often depends on the representative chosen for gH. For example, consider

$$G = \mathcal{S}_3 \text{ and } H = \langle (1\,2) \rangle = \{(1), (1\,2)\}.$$

Then we would have that $(1\,2)H \otimes (1\,2\,3)H = (2\,3)H$ and $(1)H \otimes (1\,2\,3)H = (1\,2\,3)H$. However we have that

$$(1\,2)H = \{(1), (1\,2)\} = (1)H \quad \text{but} \quad (1\,2\,3)H = \{(1\,2\,3), (1\,3)\} \neq \{(2\,3)H = \{(2\,3), (1\,3)\}.$$

What goes wrong here is that the subgroup H is not a normal subgroup of $\mathcal{S}_3$ for in general, if H is not normal, then there exists an element h which is not in the kernel of the action of $G \circlearrowright G/H$. This means there is a coset aH with

$$hH * aH = h \bullet aH \neq aH.$$

However, since $h \in H$ we have that $hH = eH = H$. Therefore, if $*$ were really well defined we would indeed have that

$$hH * aH = eH * aH = eaH = aH.$$

It is a pleasant surprise that the converse is also true. This means that if $H \trianglelefteq G$ then this crap can't happen (i.e. $*$ is indeed well-defined). This is the content of the theorem below.

Theorem 11.1. *$*$ is well-defined iff $H \trianglelefteq G$.*

Proof$\Leftrightarrow$ We have already proven this direction.

⇐ Now suppose $H \trianglelefteq G$. This means for any $h \in H$ and any coset $C \in G/H$ we have that $h \bullet C = C$. Now, what we need to show is if $aH = a'H$ and $bH = b'H$, then $abH = a'b'H$. But $aH = a'H$ means that there exists an $h \in H$ so that $a' = ah$ and $bH = b'H$ means there exists a (possibly different) element $h' \in H$ so that $b' = bh'$. Therefore

$$(a' \cdot b')H = (ah \cdot bh')H = a \bullet (h \bullet (b \bullet h'H)) = a \bullet (h \bullet bH) = a \bullet bH = (a \cdot b)H.$$

Therefore we have proved that $a'b'H = abH$ and so $*$ is well defined, as desired. ∎

Thus if $H \trianglelefteq G$ is normal in G we have a binary operation

$$\begin{aligned} *&: {}^{G}\!/_{H} \times {}^{G}\!/_{H} \to {}^{G}\!/_{H} \\ &: g_1H * g_2H = g_1g_2H. \end{aligned}$$

The obvious question is **does this give us a group?** The answer is yes. This follows from the fact that when $*$ is well defined the map

$$g \mapsto gH$$

is surjective and operation preserving.

EXERCISE 11.1. Let G be a group and let X be an algebraic system. Suppose that we an operation preserving map from G onto X. Prove that X is a group.
Hint! It turns out that the identity of X is the image of the identity of G under φ. [I]

SOLUTION

Using this exercise and the comment made before the exercise, we therefore have a new group (no, The Beetles didn't get back together)! We summarize this:

THEOREM 11.2. *Suppose that N is a normal subgroup of a group G. Then $(G/N, *)$ is a group where $(aN) * (bN) := abN$. Moreover we have*

- $(aN)^{-1} = a^{-1}N$
- $e_{G/N} = e_G N = N$. ∎

EXAMPLE 11.3. If N has index 2 in G, then G/N being a group of order 2, must be isomorphic to $\mathbb{Z}_2$. As such we have one coset C in addition to N. They follow the table:

$*$	N	C
N	N	C
C	C	N

.

In particular, for any $g_1, g_2 \in G$ which are both not in N, then both g_1N and g_2N cannot equal N, so they must equal C. Therefore

$$(g_1 \cdot g_2)N = g_1N * g_2N = C * C = N \Rightarrow g_1 \cdot g_2 \in N.$$

For instance, the set of even permutations of degree n, $\mathcal{A}_n$, are an index two subgroup of $\mathcal{S}_n$. In this case what we just proved in general generalizes the fact that the product of any two odd permutations (which make up the coset C in the above notation) is an even permutation.

Another place you've seen this is in the study of symmetry. There, this general fact corresponds to the fact that the product of any two reflections is a rotation.

Example 11.4. Consider the group $G = \mathbb{Z}$ and for a fixed positive integer n, $N = n\mathbb{Z}$. (Of course this subgroup is normal in $\mathbb{Z}$ since every subgroup of $\mathbb{Z}$ (or any abelian group for that matter) is normal.) In this case we have

$$\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}_n.$$

In fact we showed in Example 10.51 that the cosets coincide with the elements of $\mathbb{Z}_n$. Moreover, we clearly have that the operation on these cosets coincides with the operation we defined on $\mathbb{Z}_n$ so very long ago. In this way you should think of quotient groups as a generalization of the integers mod n.

Example 11.5. In $\mathcal{D}_4$, the 180° degree rotation ρ^2 is central. Therefore $N := \langle \rho^2 \rangle \trianglelefteq \mathcal{D}_4$. In the table the four different cosets of N in G are colored by four different colors. For instance, the table says that

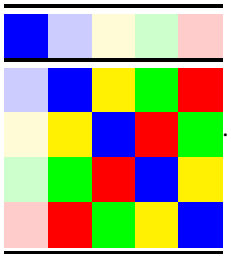
yellow $\otimes$ red = green.

$\mathcal{D}_4$	$\mathbf{1}$	ρ^2	ρ	ρ^3	$\mathbf{r}$	$\rho^2\mathbf{r}$	$\rho\mathbf{r}$	$\rho^3\mathbf{r}$
$\mathbf{1}$	$\mathbf{1}$	ρ^2	ρ	ρ^3	$\mathbf{r}$	$\rho^2\mathbf{r}$	$\rho\mathbf{r}$	$\rho^3\mathbf{r}$
ρ^2	ρ^2	$\mathbf{1}$	ρ^3	ρ	$\rho^2\mathbf{r}$	$\mathbf{r}$	$\rho^3\mathbf{r}$	$\rho\mathbf{r}$
ρ	ρ	ρ^3	ρ^2	$\mathbf{1}$	$\rho\mathbf{r}$	$\rho^3\mathbf{r}$	$\rho^2\mathbf{r}$	$\mathbf{r}$
ρ^3	ρ^3	ρ	$\mathbf{1}$	ρ^2	$\rho^3\mathbf{r}$	$\rho\mathbf{r}$	$\mathbf{r}$	$\rho^2\mathbf{r}$
$\mathbf{r}$	$\mathbf{r}$	$\rho^2\mathbf{r}$	$\rho^3\mathbf{r}$	$\rho\mathbf{r}$	$\mathbf{1}$	ρ^2	ρ^3	ρ
$\rho^2\mathbf{r}$	$\rho^2\mathbf{r}$	$\mathbf{r}$	$\rho\mathbf{r}$	$\rho^3\mathbf{r}$	ρ^2	$\mathbf{1}$	ρ	ρ^3
$\rho\mathbf{r}$	$\rho\mathbf{r}$	$\rho^3\mathbf{r}$	$\mathbf{r}$	$\rho^2\mathbf{r}$	ρ	ρ^3	$\mathbf{1}$	ρ^2
$\rho^3\mathbf{r}$	$\rho^3\mathbf{r}$	$\rho\mathbf{r}$	$\rho^2\mathbf{r}$	$\mathbf{r}$	ρ^3	ρ	ρ^2	$\mathbf{1}$

If we take the table for $\mathbb{Z}_2 \times \mathbb{Z}_2$

$\mathbb{Z}_2 \times \mathbb{Z}_2$	$([0]_2, [0]_2)$	$([1]_2, [0]_2)$	$([0]_2, [1]_2)$	$([1]_2, [1]_2)$
$([0]_2, [0]_2)$	$([0]_2, [0]_2)$	$([1]_2, [0]_2)$	$([0]_2, [1]_2)$	$([1]_2, [1]_2)$
$([1]_2, [0]_2)$	$([1]_2, [0]_2)$	$([0]_2, [0]_2)$	$([1]_2, [1]_2)$	$([0]_2, [1]_2)$
$([0]_2, [1]_2)$	$([0]_2, [1]_2)$	$([1]_2, [1]_2)$	$([0]_2, [0]_2)$	$([1]_2, [0]_2)$
$([1]_2, [1]_2)$	$([1]_2, [1]_2)$	$([0]_2, [1]_2)$	$([1]_2, [0]_2)$	$([0]_2, [0]_2)$

we see that both color patterns are exactly the same, they are



This is a visual way of proving that

$$\mathcal{D}_4/\langle \rho^2 \rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2.$$

Many facts about quotient groups are proven simply by observing that

$$\begin{aligned}\pi_N &: G \to G/N \\ &: g \mapsto gN\end{aligned}$$

is a surjective group homomorphism with kernel N (this is the definition of π_N). This is almost true by definition. Since

$$\pi_N(g \cdot \dot{g}) = g\dot{g}N = gN * \dot{g}N = \pi_N(g) * \pi_N(\dot{g})$$

we have that π_N is a homomorphism. Moreover, since N is the identity of G we compute the kernel as

$$\begin{aligned}\text{Ker}\,(\pi_N) &= \{g \in G : \pi_N(g) = N\} \\ &= \{g \in G : gN = N\} \\ &= \{g \in G : e_G^{-1} g \in N\} \\ &= \{g \in G : g \in N\} \\ &= N.\end{aligned}$$

Lastly, like all quotient maps π_N (from a set to that same set mod an equivalence relation) is surjective. In more detail, an arbitrary element of G/N has the form gN, and we $g \mapsto gN$.

For instance, translating the basic homomorphism facts given in Theorem 8.10 gives us (only the third part is new!)

THEOREM 11.6.

(1) $e_{G/N} = N$

(2) $(gN)^{-1} = g^{-1}N$

(3) $(gN)^n = g^n N$.

The above facts will be used often and without explicit reference!

THEOREM 11.7. *If N is a normal subgroup of G then*

(1) *The order of gN in the group G/N, denoted as usual as $|\,gN\,|$, is equal to the first positive integer n so that $g^n \in N$.*

(2) $G = \langle\, x_1, \ldots, x_n \,\rangle \Rightarrow G/N = \langle\, x_1 N, \ldots, x_n N \,\rangle$;

(3) *hN and kN commute iff $hkh^{-1}k^{-1} \in N$;*

(4) *Given a homomorphism $\varphi : G \to K$, the map*

$$\begin{aligned}\widetilde{\varphi} &: G/N \to K \\ &: gN \mapsto \varphi(g)\end{aligned}$$

is well defined iff $N \subseteq \text{Ker}\,(\varphi)$.

[1] Of course since $g^{|g|} = e_G \in N$, this number is at most $|\,g\,|$.

Remark 11.8. A particular case of this theorem, 11.7**(4)**, says for $N_1, N_2 \trianglelefteq G$, since the quotient map $G \to G/N_2$ is a group homomorphism with kernel N_2, (so we're using G/N_2 as K in the statement) the map

$$\begin{aligned} \pi_{N_2}^{N_1} &: G/N_1 \to G/N_2 \\ &: gN_1 \mapsto gN_2 \end{aligned}$$

is well defined (and induced from $G \to G/N_2$) iff $N_1 \preccurlyeq N_2$.

Proof.

(1) Let $\overline{g} := g^n$. Then we have that

$$(gN)^n = g^n N = \overline{g}N = N \tag{1}$$

holds iff $\overline{g} \in N$. Therefore, the first positive integer n so that (1) holds is the first positive integer n so that $g^n \in N$.

(2) This part is a direct consequence of Proposition 9.26**(1)**. You can see this by letting $\varphi = \pi_n$ and $H = G/N$. In words G/N is generated by the image of a generating set of G.

(3) Suppose that hN and kN commute. Then we have that

$$\begin{aligned} (hN) * (kN) = (kN) * (hN) &\Rightarrow (hN) * (kN) * (hN)^{-1} * (kN)^{-1} = e_{G/N} \\ &\Rightarrow hN * kN * h^{-1}N * k^{-1}N \\ &\Rightarrow \left(h \cdot k \cdot h^{-1} \cdot k^{-1}\right) N = e_{G/N} = N \\ &\Rightarrow hkh^{-1}k^{-1} \in N. \end{aligned}$$

The converse is proven by turning all of the arrows around.

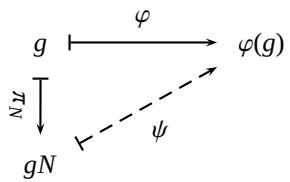
(4) We apply Lemma 8.66 back on page 420 with G/N playing the role of H, $\varphi_1 = \pi_N$, the quotient map, K starring as itself, and φ playing the role of φ_2. The lemma says that the relation $\psi := \varphi_2 \circ \varphi_1^{-1}$ is a function iff

$$\operatorname{Ker}(\varphi_1) \subseteq \operatorname{Ker}(\varphi_2).$$

Plugging in our actors and using the fact that $\operatorname{Ker}(\pi_N) = N$, this is the statement

$$N \subseteq \operatorname{Ker}(\varphi).$$

Lastly, since we have



we see that $\psi(gN) = \varphi(g)$ as stated. ∎

In general, we have three ways of computing the isomorphism type of a quotient. The second two methods are where I would like you to be, but the first method is where you should start to get a little bit used to this!

Method 1 Cut up the table for the original group into pieces according to which coset things are in. We used this method in Example 11.5 to show that $\mathcal{D}_4/\langle \rho^2 \rangle \cong \mathbb{Z}_2^2$. This method has extremely limited use. For example, it only works for small groups since tables for larger groups quickly get out of hand! Moreover you need to pay attention before hand to how you cut this up since a poor choice will not be very useful. Another disadvantage of this method is that when you are done perhaps you don't recognize what the quotient actually is! Perhaps this is because the resulting table ends up being written in a different order than you may be used to.

Method 2 Explicitly write out the multiplication table of the cosets. This method is faster than the previous since we only need to write the table for the quotient and not for the original group. This is a good way to start, but it also quickly begins to suck and suck bad. For instance, we need to list all of the cosets just to start this method, and often that is quite a pain in the rear.

Method 3 Compute the image of a (your favorite!) generating set in the quotient. Then see if we can write down some basic properties of this generating set to identify the quotient. For instance, is the quotient a group of order 8 and generated by two elements of order 2? Then we know it is isomorphic to $\mathcal{D}_4$.

Method 4 Use the classification of groups of order $\frac{|G|}{|N|}$ and then perform a process of elimination. For instance, does our quotient have order 3? Then the group must be isomorphic to $\mathbb{Z}_3$ since this is the **only** group of order 3! Is our group a non-abelian group of order 14? Then it is isomorphic to $\mathcal{D}_5$. Is it a group of order 4 with (at least) two elements of order 2? Then it's isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. For instance, our example, Example 11.5, we could have simply noted that the elements (i.e. cosets) of $\mathcal{D}_4/\langle \rho^2 \rangle$, $\rho\langle \rho^2 \rangle$ and $\mathbf{r}\langle \rho^2 \rangle$, are distinct elements of order 2 (they have order 2 since both $(\rho)^2$ and $(\mathbf{r})^2$ are in $\langle \rho^2 \rangle$ and they are distinct since their "ratio" $\mathbf{r}^{-1}\rho$ is not.

Of course, method **(4)** only works when we have classified the groups of order $|G|/|N|$! It is very efficient when we have a very short list of possibilities.

EXAMPLE 11.9. Let

$$G = B\left(\blacktriangleright^3\right) = \left\langle \mathbf{i}, \frac{1}{2} + \frac{1}{2}\mathbf{i} + \overset{\omega}{\overset{\shortparallel}{\frac{1}{2}\mathbf{i}}} + \frac{1}{2}\mathbf{k} \right\rangle.$$

We have shown that $Q_8 = \langle \mathbf{i}, \mathbf{j} \rangle$ is normal in $B\left(\blacktriangleright^3\right)$. I illustrate the latter three methods in reverse order.

- Since
$$B\left(\blacktriangleright^3\right)/Q_8$$
has $24/8 = 3$ elements, this quotient must be isomorphic to $\mathbb{Z}_3$ since this is the only group of order 3.

- The quotient $B\left(\blacktriangleright^3\right)/Q_8$ is generated by
$$\langle \mathbf{i}Q_8, \omega Q_8 \rangle = \langle \omega Q_8 \rangle.$$
We can ignore the $\mathbf{i}Q_8$ element since $\mathbf{i} \in Q_8$ and so
$$\mathbf{i}Q_8 = Q_8 = e_{B\left(\blacktriangleright^3\right)/Q_8}$$
and we do not need to write the identity down in any generating set. Therefore the quotient is cyclic, and we simply need to compute the order its generator, ωQ_8. Since
$$\omega \notin Q_8 \qquad \omega^2 = -\frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k} \notin Q_8 \qquad \omega^3 = -1 \in Q_8 \Rightarrow (\omega Q_8)^3 = Q_8$$
this order is 3. Therefore, we have that the quotient is cyclic and generated by a single element of order 3, therefore it is isomorphic to $\mathbb{Z}_3$.

- This (first) method in this case isn't too much harder than the second (the one we just did). Here we have

$$B\left(\blacktriangleright^3\right)\!/_{Q_8} = \{Q_8, \omega Q_8, \omega^2 Q_8\}.$$

If we set up the table we have

$B\left(\blacktriangleright^3\right)/Q_8$	Q_8	ωQ_8	$\omega^2 Q_8$
Q_8	Q_8	ωQ_8	$\omega^2 Q_8$
ωQ_8	ωQ_8	$\omega^2 Q_8$	Q_8
$\omega^2 Q_8$	$\omega^2 Q_8$	Q_8	ωQ_8

Of course this table (shows that the quotient) is isomorphic to $\mathbb{Z}_3$.

Example 11.10. Let's kick this up a bit by now considering

$$B\left(\blacklozenge^3\right) = \left\langle \frac{1+\mathbf{i}}{\sqrt{2}}, \frac{1+\mathbf{j}}{\sqrt{2}} \right\rangle$$

(again) with normal subgroup Q_8. We let $S_{\mathbf{i}}$ denote the first listed generator and $S_{\mathbf{j}}$ the second. To calculate the quotient, $B\left(\blacklozenge^3\right)/Q_8$, we again illustrate the three methods in reverse order.

- By Lagrange, $B\left(\blacklozenge^3\right)/Q_8$ is a group of order $48/8 = 6$. Moreover, since

$$\begin{aligned} S_{\mathbf{i}} S_{\mathbf{j}} S_{\mathbf{i}}^{-1} S_{\mathbf{j}}^{-1} &= \left(\frac{1+\mathbf{i}}{\sqrt{2}}\right)\left(\frac{1+\mathbf{j}}{\sqrt{2}}\right)\left(\frac{1-\mathbf{i}}{\sqrt{2}}\right)\left(\frac{1-\mathbf{j}}{\sqrt{2}}\right) \\ &= \left(\frac{1+\mathbf{k}}{\sqrt{2}}\right)\left(\frac{1-\mathbf{j}}{\sqrt{2}}\right) \qquad \text{(Exercise 3.51)} \\ &= \frac{1}{2} + \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k} \notin Q_8 \end{aligned}$$

the quotient is non-abelian. Therefore, since there is only one non-abelian group of order 6, namely $\mathcal{S}_3$, this quotient must be isomorphic to $\mathcal{S}_3$.

- Since

$$S_{\mathbf{i}}^2 = \mathbf{i} \in Q_8 \text{ and } S_{\mathbf{j}}^2 = \mathbf{j} \in Q_8$$

we know that both generators of the quotient $S_{\mathbf{i}}Q_8$ and $S_{\mathbf{j}}Q_8$ have order 2 there. As such, this quotient is generated by two elements of order 2 ,and it must be a dihedral group. Moreover, since the quotient has order 6 we know that the quotient must be $\mathcal{D}_3$, which of course is isomorphic to $\mathcal{S}_3$.

- This part will certainly be no fun ¨. Notice that what we called ω from the previous part is in this group and it is equal to $S_{\mathbf{i}}S_{\mathbf{j}}$ (actually

$$B\left(\blacktriangleright^3\right) \preccurlyeq B\left(\blacklozenge^3\right),$$

this can be seen geometrically since a tetrahedron can be inscribed inside of a cube). I rerecord the first 3 powers of ω:

$$\begin{aligned} \omega &= \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k} \\ \omega^2 &= -\frac{1}{2} + \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k} \\ \omega^3 &= \frac{1}{2} + \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}. \end{aligned}$$

We can check that the 6 cosets

$$Q_8 \qquad \omega Q_8 \qquad \omega^2 Q_8 \qquad S_{\mathbf{i}} Q_8 \qquad S_{\mathbf{i}} \omega Q_8 \qquad S_{\mathbf{i}} \omega^2 Q_8$$

are distinct, and hence make up all 6 cosets of the quotient. To save our sanity, we use as many tricks as possible. Using our facts about $B\left(\blacktriangleright^3\right)$ we see right away that the first three cosets are distinct. Similarly, we can cancel out $\mathbb{S}_{\mathbf{i}}$ to see the next three cosets are also distinct from each other. Lastly the last three cosets are each distinct from each of the first three cosets since each of the last three cosets have order 2 while none of the first three have order 3.

Next, let's check out what the product

$$\omega Q_8 S_{\mathbf{i}} Q_8$$

is. Motivated by the fact that we already know the answer, we can cheat to save time and check right away that

$$\omega Q_8 S_{\mathbf{i}} Q_8 = S_{\mathbf{i}} \omega^2 Q_8$$

by, as usual, showing that

$$\left(S_{\mathbf{i}} \omega^2\right)^{-1} \omega S_{\mathbf{i}} \in Q_8.$$

Indeed,

$$\begin{aligned}
\left(S_{\mathbf{i}} \omega^2\right)^{-1} \omega S_{\mathbf{i}} &= \omega^4 S_{\mathbf{i}}^{-1} \omega S_{\mathbf{i}} \\
&= \omega^4 \cancel{S_{\mathbf{i}}^{-1}} \cancel{S_{\mathbf{i}}} S_{\mathbf{j}} S_{\mathbf{i}} \\
&= -\omega S_{\mathbf{j}} S_{\mathbf{i}} \\
&= -S_{\mathbf{i}} S_{\mathbf{j}} S_{\mathbf{j}} S_{\mathbf{i}} \\
&= -S_{\mathbf{i}} \mathbf{j} S_{\mathbf{i}} \\
&= -S_{\mathbf{i}} S_{\mathbf{i}}^{-1} \mathbf{j} \\
&= -\mathbf{j}
\end{aligned}$$

which indeed is in Q_8. Therefore, we can now make the table for the quotient quickly:

$B\left(\blacklozenge^3\right)/Q_8$	Q_8	ωQ_8	$\omega^2 Q_8$	$S_{\mathbf{i}} Q_8$	$S_{\mathbf{i}} \omega Q_8$	$S_{\mathbf{i}} \omega^2 Q_8$
Q_8	Q_8	ωQ_8	$\omega^2 Q_8$	$S_{\mathbf{i}} Q_8$	$S_{\mathbf{i}} \omega Q_8$	$S_{\mathbf{i}} \omega^2 Q_8$
ωQ_8	ωQ_8	$\omega^2 Q_8$	$\omega^2 Q_8$	$S_{\mathbf{i}} \omega^2 Q_8$	$S_{\mathbf{i}} Q_8$	$S_{\mathbf{i}} \omega Q_8$
$\omega^2 Q_8$	$\omega^2 Q_8$	Q_8	ωQ_8	$S_{\mathbf{i}} \omega Q_8$	$S_{\mathbf{i}} \omega^2 Q_8$	$S_{\mathbf{i}} Q_8$
$S_{\mathbf{i}} Q_8$	$S_{\mathbf{i}} Q_8$	$S_{\mathbf{i}} \omega Q_8$	$S_{\mathbf{i}} \omega^2 Q_8$	Q_8	ωQ_8	$\omega^2 Q_8$
$S_{\mathbf{i}} \omega Q_8$	$S_{\mathbf{i}} \omega Q_8$	$S_{\mathbf{i}} \omega^2 Q_8$	$S_{\mathbf{i}} Q_8$	$\omega^2 Q_8$	Q_8	ωQ_8
$S_{\mathbf{i}} \omega^2 Q_8$	$S_{\mathbf{i}} \omega^2 Q_8$	$S_{\mathbf{i}} Q_8$	$S_{\mathbf{i}} \omega Q_8$	ωQ_8	$\omega^2 Q_8$	Q_8

EXAMPLE 11.11. In this example, we find a group, denoted ${}_4G$, which has three normal subgroups of order 2 (which are therefore each isomorphic to $\mathbb{Z}_2$) with each of the three quotients non-isomorphic to each other. The group in question is

$${}_4G = \left\langle (1\,\overset{x}{2}\,3\,4), (1\,3)(\overset{y}{5}\,6\,7\,8) \right\rangle.$$

Let's record some basic identities that occur in this group (using these identities will, most of the time, be easier than going back to the definition of the elements x and y in terms of permutations). For starters, we have that x and y have order 4. Moreover we also have that $yxy^{-1} = x^{-1}$. Therefore

$$\begin{aligned} y^2xy^{-2} &= yx^{-1}y^{-1} \\ &= \left(yxy^{-1}\right)^{-1} \\ &= x^{-1^{-1}} \\ &= x \end{aligned}$$

and

$$\begin{aligned} yx^2y^{-1} &= \left(yxy^{-1}\right)^2 \\ &= \left(x^{-1}\right)^2 \\ &= x^{-2} \\ &= x^2. \end{aligned}$$

Therefore y^2 commutes with x and (obviously) y and the same can be said for x^2. Hence $Z := \left\langle x^2, y^2 \right\rangle \leqslant \mathcal{C}(G4)$. (Actually, with a bit more work we can see that it is the entire center, but we will get this for free from Exercise 11.35.**??** coming up on page **??**. Since Z is a subgroup of the center of $G4$ it is normal in $G4$. Moreover, we obviously have that

$$Z = \{(1), (1\,3)(2\,4), (5\,7)(6\,8), (1\,3)(2\,4)(5\,7)(6\,8)\}.$$

In particular we note that no element of Z contains a $\underline{4}$-cycle. Now since Z is normal, ${}_4G/Z$ is a group and

$${}^4G\!/_Z = \langle\, xZ, yZ \,\rangle.$$

I claim that $x' := xZ$ and $y' := yZ$ are also 2 distinct elements of order 2 which commute (this will establish, ⟨by coincidence ⟩ the quotient has ⟨also ⟩ order four ⟨moreover it, like Z itself, is also isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. ⟩)

- We first show that x' and y' have order 2. For starters, x^2 and y^2 are in Z by definition of Z. Hence $(x')^2 = (y')^2 = e_{G/Z}$. Therefore to see that each have order 2 we simply need to see that they are not the identity. That is we need to show that x and y are not in Z. But both x and y contain $\underline{4}$-cycles and, as such, are not in Z.
- Next we need to check that x' and y' are not equal. But x' and y' are equal iff $y^{-1}x \in Z$. But $y^{-1}x$ also contains a $\underline{4}$-cycle, and we therefore have that $x' \neq y'$.
- The last thing we need to check is that x' and y' commute. But

  $$yxy^{-1}x^{-1} = x^{-1}x^{-1} = x^{-2} = x^2 \in Z.$$

  Therefore

  $$y'x'y'^{-1}x'^{-1} = e_{G/Z} \Rightarrow x'y' = y'x'$$

  as desired.

In conclusion, by Lagrange's theorem we have that

$$|\,{}_4G\,| = |\,Z\,|\left|\,{}^4G\!/_Z\,\right| = 4 \cdot 4 = 16.$$

Example 11.12. Let ${}_4G$ be as in the previous example and define

$$N_1 := \left\langle x^2 \right\rangle = \{e, x^2\} \qquad N_2 := \left\langle x^2y^2 \right\rangle = \{e, x^2y^2\} \quad \text{and} \quad N_3 := \left\langle y^2 \right\rangle = \{e, y^2\}$$

Each of these three groups have order 2 and are in the center. In particular, each are normal subgroups of ${}_4G$. We try to find the three quotients.

(1) In

$${}_4G/N_1 = \left\langle \overset{x'}{\overset{\shortparallel}{xN_1}}, \overset{y'}{\overset{\shortparallel}{yN_1}} \right\rangle$$

we that $|x'| = 2, |y'| = 4$ (the latter follows from the fact that we have $\langle x \rangle \cap \langle y \rangle$ is trivial since every power of y moves the letter $5 \in \overline{8}$ but no power of x moves 5). Next note that $\langle x' \rangle \cap \langle y' \rangle$ is also trivial since we must have that $(y')^2 \neq x'$ since the latter is a subset of $\langle x \rangle$. Lastly we also have that y' and x' commute since

$$yxy^{-1}x^{-1} = x^3x^{-1} = x^2 \in H_1.$$

Thus ${}_4G/H_1 \cong \mathbb{Z}_4 \times \mathbb{Z}_2$.

(2) In ${}_4G/N_2$ we again write x' for the image of x and y' for the image of y. This time, however y' is the one of order 2 while x' has order 4. We also have that $y', x'y'$ also generates ${}_4G/N_2$ and

$$(x'y')^2 = x'y'x'y' = x'y'x'y'^{-1} = x'(x')^{-1} = e$$

implies that both have order 2. Lastly, since $(x'y')y' = x'$ has order 4, the quotient must be $\mathcal{D}_4$ by Theorem **??**.

(3) Possibly the most interesting example comes from considering ${}_4G/N_3$. We always we let x' and y' have their usual meanings. In this case we can check directly that they are distinct elements, each of order 4 (actually we can see they do not commute, so they are obviously distinct). However we also have that

$$x^2y^2 \in N_3 \Rightarrow x'^2 = y'^{-2} = y'^2.$$

Lastly, in this final example we have that

$$y'x'y'^{-1}x'^{-1} = x'^2 \neq e_{{}_4G/N_3}$$

since $x^2 = x^2y^2$ implies that $y^2 = e_{{}_4G}$ which is false since y has order 4. Hence the quotient is generated by 2 non-commuting elements of order 4 and is therefore isomorphic to Q_8.

Exercise 11.2. I hope this exercise eliminates the natural tendency of the student to say "oh, it looks like division, so I can apply all the properties of division of numbers!"

Solution

(1) Suppose that N_1 and N_2 are two isomorphic normal subgroups of a group G. Is it true that G/N_1 is isomorphic to G/N_2?

(2) Suppose that N_1 and isomorphic to N_2 but G_1 is not isomorphic to G_2. Is it true that G_1/N_1 is not isomorphic to G_2/N_2?

(3) Suppose that N_1 and N_2 are normal subgroups of a group G. Suppose further that G/N_1 is isomorphic to G/N_2. Is it true that N_1 is isomorphic to N_2?

Hint! All three are false!!

11.2 Isomorphism Theorems

Theorem 11.13 (The First Isomorphism Theorem). *Let φ be a homomorphism from G to H. Then we have*

$$G/\mathrm{Ker}(\varphi) \cong \mathrm{Im}(\varphi).$$

Proof. We have an epimorphism from G to $G/\mathrm{Ker}(\varphi)$ and from G to $\mathrm{Im}(\varphi)$. Both of these epimorphisms have kernel equal to $\mathrm{Ker}(\varphi)$. Therefore, by Theorem 8.65 both are isomorphic. ∎

We can think of this theorem of being a three step process:

- Define a homomorphism $\varphi: G \to H$.
- Find the image of φ.
- Find the kernel of φ.
- Conclude there is isomorphism which follows the first isomorphism theorem.

Of course we have been doing the first three steps of this for some time now!

Actually, using the lemma from that section, Lemma 8.66 we have the following

Proposition 11.14. *Suppose that $\varphi: G \to \widehat{G}$ is a homomorphism of groups, and $N \trianglelefteq G$. Then $\widetilde{\varphi}(gN) := \varphi(g)$ is well defined iff $N \subseteq \mathrm{Ker}(\varphi)$.*

Proof. Use Lemma 8.66 with $\varphi_1 = \pi_N$, the quotient map $G \to G/N$, $\varphi_2 = \varphi$, the given map. Since $\mathrm{Ker}(\pi_N) = N$, we have that $\mathrm{Ker}(\pi_N) \preccurlyeq \mathrm{Ker}(\varphi)$ holds iff $N \subseteq \mathrm{Ker}(\varphi)$. ∎

Example 11.15.

(1) There exists an epimorphism from $\mathbb{Z}$ to $\mathbb{Z}_n$ given by $m \mapsto [\,m\,]_n$. The kernel is $n\mathbb{Z}$. Therefore we recover our earlier observation that

$$\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n.$$

(2) There exists an epimorphism

$$\begin{aligned} \exp^*: \mathbb{R} &\to \mathbb{S}^1 \\ : t &\mapsto e^{2\pi i t}. \end{aligned}$$

The kernel of this homomorphism is the set of integers (i.e. $\mathrm{Ker}(\exp^*) = \mathbb{Z}$). This is because

$$\exp^*(t) = e^{2\pi i t} = \cos 2\pi t + \sin 2\pi t\, i = 1$$

iff

$$\cos 2\pi t = 1 \quad \text{and} \quad \sin 2\pi t = 0$$

which is true iff $t \in \mathbb{Z}$.

Therefore we have that

$$\mathbb{R}/\mathbb{Z} \cong \mathbb{S}^1.$$

Actually, it is not too difficult to see this isomorphism by thinking of $\mathbb{R}$ as a helix over the circle.

(3) Consider the function

$$\begin{aligned}\varphi\colon \mathrm{UT}_3\left(()\,3,\mathbb{Z}_2\right) &\to \mathbb{Z}_2\times\mathbb{Z}_2\\ :\begin{pmatrix}[\,1\,]_2 & [\,a\,]_2 & [\,b\,]_2\\ [\,0\,]_2 & [\,1\,]_2 & [\,c\,]_2\\ [\,0\,]_2 & [\,0\,]_2 & [\,1\,]_2\end{pmatrix} &\mapsto ([\,a\,]_2\,,[\,b\,]_2).\end{aligned}$$

Then

(a) Let's do two calculations and compare them. Let

$$\mathbf{A}:=\begin{pmatrix}[\,1\,]_2 & [\,a\,]_2 & [\,b\,]_2\\ [\,0\,]_2 & [\,1\,]_2 & [\,c\,]_2\\ [\,0\,]_2 & [\,0\,]_2 & [\,1\,]_2\end{pmatrix}\quad\text{and}\quad \mathbf{A}':=\begin{pmatrix}[\,1\,]_2 & [\,a'\,]_2 & [\,b'\,]_2\\ [\,0\,]_2 & [\,1\,]_2 & [\,c'\,]_2\\ [\,0\,]_2 & [\,0\,]_2 & [\,1\,]_2\end{pmatrix}$$

be two arbitrary matrices in the group $\mathrm{UT}_3\left(\mathbb{Z}_2\right)$. Then we have

$$\begin{aligned}\mathbf{A}\cdot\mathbf{A}' &= \begin{pmatrix}[\,1\,]_2 & [\,a\,]_2 & [\,b\,]_2\\ [\,0\,]_2 & [\,1\,]_2 & [\,c\,]_2\\ [\,0\,]_2 & [\,0\,]_2 & [\,1\,]_2\end{pmatrix}\cdot\begin{pmatrix}[\,1\,]_2 & [\,a'\,]_2 & [\,b'\,]_2\\ [\,0\,]_2 & [\,1\,]_2 & [\,c'\,]_2\\ [\,0\,]_2 & [\,0\,]_2 & [\,1\,]_2\end{pmatrix}\\ &= \begin{pmatrix}[\,1\,]_2 & [\,a\,]_2+[\,a'\,]_2 & [\,b'\,]_2+[\,ac'\,]_2+[\,b\,]_2\\ [\,0\,]_2 & [\,1\,]_2 & [\,c\,]_2+[\,c'\,]_2\\ [\,0\,]_2 & [\,0\,]_2 & [\,1\,]_2\end{pmatrix}.\end{aligned}$$

Therefore, we have that

$$\varphi(\mathbf{A}\cdot\mathbf{A}')=([\,a\,]_2+[\,a'\,]_2\,,[\,b\,]_2+[\,b'\,]_2).$$

Meanwhile, we have

$$\begin{aligned}\varphi(\mathbf{A})+\varphi(\mathbf{A}') &= \varphi\left(\begin{pmatrix}[\,1\,]_2 & [\,a\,]_2 & [\,b\,]_2\\ [\,0\,]_2 & [\,1\,]_2 & [\,c\,]_2\\ [\,0\,]_2 & [\,0\,]_2 & [\,1\,]_2\end{pmatrix}\right)+\varphi\left(\begin{pmatrix}[\,1\,]_2 & [\,a\,]_2 & [\,b\,]_2\\ [\,0\,]_2 & [\,1\,]_2 & [\,c\,]_2\\ [\,0\,]_2 & [\,0\,]_2 & [\,1\,]_2\end{pmatrix}\right)\\ &= ([\,a\,]_2\,,[\,b\,]_2)+([\,a'\,]_2\,,[\,b'\,]_2)=([\,a\,]_2+[\,a'\,]_2\,,[\,b\,]_2+[\,b'\,]_2).\end{aligned}$$

In conclusion $\varphi(\mathbf{A}\cdot\mathbf{A}')=\varphi(\mathbf{A})+\varphi(\mathbf{A}')$ and, by definition, φ is a homomorphism of groups.

(b) The homomorphism φ is clearly surjective. That is the image of φ is all of $\mathbb{Z}_2\times\mathbb{Z}_2$. This is because for any $([\,a\,]_2\,,[\,b\,]_2)\in\mathbb{Z}_2\times\mathbb{Z}_2$ the matrix

$$\begin{pmatrix}[\,1\,]_2 & [\,a\,]_2 & [\,0\,]_2\\ [\,0\,]_2 & [\,1\,]_2 & [\,b\,]_2\\ [\,0\,]_2 & [\,0\,]_2 & [\,1\,]_2\end{pmatrix}$$

maps to it.

(c) We now calculate the kernel:

$$([a]_2, [c]_2) = \varphi\left(\begin{pmatrix} [1]_2 & [a]_2 & [b]_2 \\ [0]_2 & [1]_2 & [c]_2 \\ [0]_2 & [0]_2 & [1]_2 \end{pmatrix}\right) = ([0]_2, [0]_2) \Rightarrow [a]_2 = 0, [b]_2 = 0$$

$$\Rightarrow \begin{pmatrix} [1]_2 & [a]_2 & [b]_2 \\ [0]_2 & [1]_2 & [c]_2 \\ [0]_2 & [0]_2 & [1]_2 \end{pmatrix} = \begin{pmatrix} [1]_2 & [0]_2 & [b]_2 \\ [0]_2 & [1]_2 & [0]_2 \\ [0]_2 & [0]_2 & [1]_2 \end{pmatrix}.$$

In other words the kernel, which we denote $\mathcal{K}$, is the subset of $\mathrm{UT}_3(\mathbb{Z}_2)$ given by

$$\left\{ \begin{pmatrix} [1]_2 & [0]_2 & [b]_2 \\ [0]_2 & [1]_2 & [0]_2 \\ [0]_2 & [0]_2 & [1]_2 \end{pmatrix} : [b]_2 \in \mathbb{Z}_2 \right\}$$

i.e., those matrices whose only non-zero entry (besides the diagonal) lies in the upper right hand corner. Note that there are only two elements of this group and this subgroup is isomorphic to $\mathbb{Z}_2$.

(d) We conclude that

$$\mathrm{UT}_3(\mathbb{Z}_2)/\mathcal{K} \cong \mathbb{Z}_2 \times \mathbb{Z}_2.$$

REMARK 11.16. Note that part **(3)** still works if we replace $\mathbb{Z}_2$ with any other ring of coefficients. For instance, we would have an isomorphism

$$\mathrm{UT}_3(\mathbb{Z}_{17})/\mathcal{K} \cong \mathbb{Z}_{17} \times \mathbb{Z}_{17}.$$

where $\mathcal{K}$ is defined in a similar way; as the set 3 by 3 matrices with 1's down the diagonal and 0's in all sother entries except for the upper right hand entry.

We given another proof of **KIT**, Theorem 10.107 back on page 560.

COROLLARY 11.17 (=Theorem 10.107)**.** *For any homomorphism φ with domain G we have*

$$|G| = |\mathrm{Ker}(\varphi)| \cdot |\mathrm{Im}(\varphi)|.$$

In particular, the order of the image of any homomorphism from G divides the order of G.

PROOF. This formula is simply a combined use of the first isomorphism theorem and Lagrange's theorem. In more detail: The first isomorphism theorem says that we have a isomorphism of groups

$$G/\mathrm{Ker}(\varphi) \cong \mathrm{Im}(\varphi).$$

By Lagrange's theorem we know that

$$\#\left(G/\mathrm{Ker}(\varphi)\right) = \frac{\#G}{\#\mathrm{Ker}(\varphi)}$$

and so[I]

$$\frac{|G|}{|\mathrm{Ker}(\varphi)|} = |\mathrm{Im}(\varphi)| \Rightarrow |G| = |\mathrm{Im}(\varphi)| \cdot |\mathrm{Ker}(\varphi)|,$$

which is the formula we wanted to prove. ∎

[I] Remember that for a group G, $\#G = |G|$. The main reason I've made the distinction is because it gives me an excuse to remind you as much as possible what is and what is not a group!

Remark 11.18. We have proven already proven this corollary in Theorem 10.107.

Remark 11.19. Note that a group H is (isomorphic to) a subgroup of G iff there exists an injective homomorphism $H \to G$. Indeed, if H were already a subgroup of G, then the inclusion map ι_H^K is an injective homomorphism. One consequence of the first isomorphism theorem just proven is that H is (isomorphic to) a quotient group of G if there exists an epimorphism $G \to H$.

One could now go back over every homomorphism we have talked about in these notes and come up with an isomorphism. For instance, we have that $\mathcal{S}_4/V \cong \mathcal{S}_3$ since we have seen an epimorphism from $\mathcal{S}_4$ to $\mathcal{S}_3$ with kernel V. Or we have that $\mathbf{GL}_n(R)/\mathbf{SL}_n(R) \cong \mathcal{U}(R)$ since again we have an epimorphism (given by the determinant) from $\mathbf{GL}_n(R)$ to $\mathcal{U}(R)$ which has kernel $\mathbf{SL}_n(R)$.

Definition 11.20. An element r in a ring R is a *perfect square* if there exists an element $a \in R$ with $a^2 = r$.[II]

Exercise 11.3.

Solution

(1) For p an odd prime, prove that half of all non-zero elements of $\mathbb{Z}_p$ are perfect squares (and the other half of non-zero elements are not.)

$$\left\langle \text{Therefore there are } (p-1)/2 \text{ perfect squares in } \mathbb{Z}_p. \right\rangle$$

Hint! Define the function

$$\begin{aligned} \hat{}^2 : \mathbb{Z}_p^\times &\to \mathbb{Z}_p^\times \\ : a &\mapsto a^2. \end{aligned}$$

Prove that $\hat{}^2$ is a homomorphism with kernel ± 1[III] and image the set of perfect squares. Then use **KIT**.

(2) More generally, prove for any commutative ring R that

$$|\mathcal{U}(R)| = \#\{r \in R : r^2 = 1\} \cdot \#\{\text{Invertible Perfect Squares in } R\} = \mathbf{O}_2(\mathcal{U}(R)) \cdot \left|\mathcal{U}(R)^{\times 2}\right|.$$

(3) How many invertible perfect squares are there in $\mathbb{Z}_{7875}$?
Hint! Use

$$7875 = 3^2 \cdot 5^3 \cdot 7 \Rightarrow \mathcal{U}(\mathbb{Z}_{7875}) \cong \mathcal{U}(\mathbb{Z}_{3^2}) \times \mathcal{U}(\mathbb{Z}_{5^3}) \times \mathcal{U}(\mathbb{Z}_7) \cong \mathbb{Z}_8 \times \mathbb{Z}_{124} \times \mathbb{Z}_6.$$

to be able to quickly compute $\mathbf{O}_2(\mathcal{U}(\mathbb{Z}_{7875}))$.

(4) How many invertible perfect squares are there in $\mathbb{Z}_{5280}$?

(5) List the 8 perfect squares in $\mathbb{Z}_{17}$.

[II] For instance, every positive real number is a perfect square and *every* complex number is a perfect square. For the rational numbers things are not as easy to state.

Example 11.21. In particular, this shows that $\mathbb{Z}_p^\times/\mathbb{Z}_p^{\times 2} \cong \mathbb{Z}_2$. As such, any two non-perfect squares in $\mathcal{U}(\mathbb{Z}_p) = \mathbb{Z}_p^\times$ have a product which *is* a perfect square. This is certainly not true for say $\mathbb{Q}^\times$: for instance neither 2 nor 3 are perfect squares (in $\mathbb{Q}$), but neither is their product. See Exercise 11.4. Also rather interesting is when -1 is not a perfect square (i.e. for $p \equiv_4 3$) then exactly one of $\pm[a]_p$ is a perfect square for every $[a]_p \neq [0]_p$.

For instance, in $\mathbb{Z}_{11}$ the perfect squares are

$$1^2 = 1 \qquad 2^2 = 4 \qquad 3^2 = 9 \qquad 4^2 \equiv_{11} 5 \qquad 5^2 \equiv_{11} 3^{\text{IV}}$$

and so the non-perfect squares (mod 11) are what's left over (or just take negatives of all the perfect squares we already found)

$$10 \qquad 7 \qquad 2 \qquad 6 \qquad 8.$$

Note that the table is

·	2	6	10	7	8
2	4	1	9	3	5
6	1	3	5	9	4
10	9	5	1	4	3
7	3	9	4	5	1
8	5	4	3	1	9

Exercise 11.4.

False Theorem 11.22. *The number* 6 *is a perfect square.*

"**Proof**". The homomorphism

$$\begin{aligned} \hat{}^2 : \mathbb{Q}^\times &\to \mathbb{Q}^\times \\ : x &\mapsto x^2 \end{aligned}$$

has kernel consisting of only ± 1 and image consisting of all perfect squares, $\mathbb{Q}^{\times 2}$. Therefore, by **KIT** we have

$$\left|\mathbb{Q}^\times/_{\{\pm 1\}}\right| = \frac{|\mathbb{Q}^\times|}{|\{\pm 1\}|} = \left|\mathbb{Q}^{\times 2}\right| \Rightarrow \left|\mathbb{Q}^\times/_{\mathbb{Q}^{\times 2}}\right| = \frac{|\mathbb{Q}^\times|}{\left|\mathbb{Q}^{\times 2}\right|} = |\pm 1| = 2.$$

Therefore $\mathbb{Q}^{\times 2}$ has index two in $\mathbb{Q}^\times$. In particular, since neither 2 nor 3 are perfect squares in $\mathbb{Q}^\times$ (hence neither are in $\mathbb{Q}^{\times 2}$) their product, which is 6, must be in in $\mathbb{Q}^{\times 2}$, c.f. Example 11.3. ■

Hint! There are no group theory mistakes. The mistake made here is quite possibly the most common error made in all of mathematics! If you're still having trouble try to isolate the first line which can be checked directly to be false–then the error must be just above!

SOLUTION

[III] This is easy, but it does take proving! For instance, it's false in $\mathbb{Z}_{15}$ since $4^2 = 1$ and $15^2 = 1$ in that ring.

[IV] Don't tell me that I'm the only one who sees $\pi = 3.14159\ldots$ here!

The following exercise illustrates how calculating the quotient is often the fastest way to compute the order of a group.

Exercise 11.5. Show that the group

$$\langle (1\,2\,3), (1\,4\,7\,2\,5\,8\,3\,6\,9) \rangle$$

has order 81.

Solution

Hint! If we call the group in question (what else?) G, one possible way of going about this is as follows:

- Show that $(4\,5\,6)$ and $(7\,8\,9)$ are in G.
- Show that $N := \langle (1\,2\,3), (4\,5\,6), (7\,8\,9) \rangle$ is a subgroup of G.
- Show that moreover $N \trianglelefteq G$.
- Show that the order of N is 27.
- Show that the 9–cycle $(1\,4\,7\,2\,5\,8\,3\,6\,9)$ is not in N (can see every element of N is a product of disjoint 3–cycles) but its cube is.
- Use the previous bullet to show that G/N has order 3.
- Use Lagrange's theorem to conclude the result.

Maple Exercise 11.6. Compute the order of the group

$$G := \langle (1\,2)(4\,5), (2\,3)(5\,6), (1\,2\,3) \rangle$$

in `GAP` or `MAPLE`. Then verify your calculation by hand.

Solution

Hint! There are a few ways of doing this. One is by showing that $(1\,2\,3)$ generates a normal subgroup of G. Then show that the quotient is isomorphic to $\mathcal{S}_3$. Yet another way is to show that the 3–cycle $(4\,5\,6)$ is in G and that $(1\,2\,3)$ and $(4\,5\,6)$ together generate a normal subgroup of order 9. Then show this quotient has order 2.

Exercise 11.7. Define the

$$\mathrm{Sgn}\colon \mathcal{S}_n \to \mathbb{Z}_2$$

as the usual ε after making the obvious isomorphism between $\mathbb{Z}_2$ and $\{\pm 1\}$ (thus Sgn of an even permutation is $[\,0\,]_2$ and Sgn of an odd permutation is $[\,1\,]_2$). Now consider

$$G := \{(\sigma, [\,n\,]_4) \in \mathcal{S}_3 \times \mathbb{Z}_4 : \mathrm{Sgn}(\sigma) = [\,n\,]_2\}$$

Show that G is a group and find its order. Which familiar group is G isomorphic to?

Solution

Hint! To compute its order consider the homomorphism from G to $\mathcal{S}_3$ defined by

$$(\sigma, [\,n\,]_4) \mapsto \sigma.$$

Show that this homomorphism is surjective and compute its kernel.)

We generalize the previous exercise, Exercise 11.7.

Exercise 11.8. In this problem, we explore yet another way of building new groups from old called the *fiber product*. The relevant ingredients are three groups: G_1, G_2 and H, and two epimorphisms $\varphi_1 : G_1 \to H$ and $\varphi_2 : G_1 \to H$. For ease of notation, we denote the pair (φ_1, φ_2) by φ. Define

$$G_1 \times_\varphi G_2 \stackrel{\text{def}}{=} \{(g_1, g_2) \in G_1 \times G_2 : \varphi_1(g_1) = \varphi_2(g_2)\}.$$

$G_1 \times_\varphi G_2$ is called the *fiber product* of G_1 and G_2 over φ.

(1) Show that $G_1 \times_\varphi G_2 \leqslant G_1 \times G_2$ (in particular, it's a group!).

(2) Define $\pi_2 : G_1 \times G_2 \to G_2$ as the standard projection, $\pi_2(g_1, g_2) = g_2$. Then denote the restriction of π_2 to the fiber product as $\pi_2|$. That is

$$\pi_2| \stackrel{\text{def}}{=} \pi_2 \,|_{G_1 \times_\varphi G_2} : G_1 \times_\varphi G_2 \to G_2.$$

(3) Prove that $\pi_2|$ is surjective.
Hint! Here, for the first time, you **must** use the fact that φ_1 is assumed to be surjective.

(4) Prove that $\mathrm{Ker}(\pi_2|) \cong \mathrm{Ker}(\varphi_1)$.
Hint! Prove/Use the fact that

$$\mathrm{Ker}(\pi_2|) = \{(g, e_{G_2}) : g \in G_1, \varphi_1(g) = e_H\}.$$

(5) Use Lagrange's theorem and the first isomorphism theorem to show that

$$\left|G_1 \times_\varphi G_2\right| = \frac{|G_1| \cdot |G_2|}{|H|}.$$

SOLUTION

Here are a few more exercises about fiber products.

Exercise 11.9.

(1) Prove that $G_1 \times_\varphi G_2 \trianglelefteq G_1 \times G_2$ iff H is abelian.

(2) Prove that $G_1 \times_\varphi G_2 = G_1 \times G_2$ iff both φ_1 and φ_2 are trivial (i.e. send every element of their domains to e_H).

(3) Prove that

$$\mathrm{Ker}(\varphi_1) \times \mathrm{Ker}(\varphi_2) \trianglelefteq G_1 \times_\varphi G_2.$$

Careful! So there are two things to show: you need to prove that $\mathrm{Ker}(\varphi_1) \times \mathrm{Ker}(\varphi_2)$ is contained inside of the fiber product **and** that $\mathrm{Ker}(\varphi_1) \times \mathrm{Ker}(\varphi_2)$ is closed under conjugation by elements of the fiber product.

SOLUTION

Exercise 11.10. Let G be a group and suppose that $z \in \mathcal{C}G$ has order n. Prove that

$$G \times \mathbb{Z}_n / \langle (z, [1]_n) \rangle \cong G.$$

Solution

Hint! The goal is to find a map $G \times \mathbb{Z}_n \to G$ with kernel $\langle (z, [1]_n) \rangle$.

Example 11.23. We can use this last exercise, Exercise 11.10, to conclude that the group from Example 11.11 which was given by

$${}_4G = \left\langle (1\ \overset{x}{2}\ 3\ 4), (1\ 3)(\overset{y}{5}\ 6\ 7\ 8) \right\rangle.$$

is not isomorphic to $G_8 \times \mathbb{Z}_2$ for any group of order 8. For suppose it were isomorphic to $G_8 \times \mathbb{Z}_2$ for some group G_8. Then as a 2-group G_8 would contain a central element of order 2, call it z and both quotients

$$G_8 \times \mathbb{Z}_2 / \langle (z, [1]_2) \rangle \quad \text{and} \quad G_8 \times \mathbb{Z}_2 / \langle (e_G, [1]_2) \rangle$$

would both be isomorphic to G_8 (here we used the previous exercise for the first isomorphism, the second isomorphism is trivial). In particular, we'd have two distinct central elements of order 2 in ${}_4G$ with isomorphic quotients. But in the example where ${}_4G$ was introduced (just referenced) we ended up showing that all such quotients were not isomorphic. (We give a more straight forward proof of the fact that ${}_4G$ and $Q_8 \times \mathbb{Z}_2$ are not isomorphic in Example 11.75 coming up on page 641).

Theorem 11.24 (The Second Isomorphism Theorem)**.** *Suppose that*

Recall that if $N \trianglelefteq G$ and $H \leqslant G$, then $H \cap N \trianglelefteq H$ and HN is a group which happens to equal NH. Note that if $H \leqslant G$, then as the image of H under the homomorphism $\pi_N : G \to G/N$, which is

$$\begin{aligned} \operatorname{Im}(\pi_N) &= \{\pi_N(h) : h \in H\} \\ &= \{hN : h \in H\}, \end{aligned}$$

would be a subgroup of G/N. Note that if $N \leqslant H$, then this is the same as H/N since in this case a coset hN is the same whether we regard things as being in G or in H. However, this is not the case if H does not contain N. Indeed, if we tack on elements of N to the end of an element of H, as in hn, since $hnN = hN$, we end up with the same image.

Theorem 11.25 (Second Isomorphism Theorem)**.** *Suppose $N, G \leqslant G$ with $N \trianglelefteq G$. Then*

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

and either can be identified with $\pi_N(H) = \{hN : h \in H\}$.

Proof. To prove the equality

$$H/N = HN/N,$$

note that since $H \leqslant HN$, we have that

$$H/N \subseteq HN/N.$$

On the other hand, every coset in HN/N has the form gN where $g \in HN$. That is $g = hn$ for some $h \in H$ and $n \in N$. But $hnN = hN \in H/N$.

Moving on, we consider the restricted quotient map

$$(\pi_N)|_H{}^{\text{V}} : H \to H/N,$$

which as always, takes $h \mapsto hN$. This map is clearly surjective, and by what we've just shown, has image HN/N.

Lastly, we need to compute the kernel of the restricted quotient map: π.

$$\begin{aligned}\text{Ker}(\pi) &= \{h \in H : \pi(h) = e_{G/N}\} \\ &= \{h \in H : \pi(h) = N\} \\ &= \{h \in H : hN = N\} \\ &= \{h \in H : h \in N\} \\ &= H \cap N.\end{aligned}$$

Therefore π is a surjective homomorphism

$$H \to HN/N$$

with kernel $H \cap N$, and so the first isomorphism theorem applies to say that

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

as desired. ∎

THEOREM 11.26 (Third Isomorphism Theorem)**.** *If $N_1 \preccurlyeq N_2 \trianglelefteq G$ and both are normal in G, then*

$$G/N_2 \cong (G/N_1)/(N_2/N_1).$$

REMARK 11.27. This isn't too difficult to remember. Simply imagine the N_1's canceling out on the right hand side to get the left hand side. This is a great example of a theorem whose proof helps understand the first isomorphism theorem better. I recall that when I first learned the first isomorphism theorem as an undergraduate I thought, "this theorem seems pretty tough, I'm not even going to look at the its secondary versions until I understand the first one better." But this was the wrong approach; it wasn't until I actually looked over the proofs of the other isomorphism theorems that I had a decent understanding of the first.

PROOF. The quotient map $G \to G/N_2$ is surjective with kernel N_2. Since $N_1 \preccurlyeq N_2$ the map

$$\begin{aligned}\pi_1^2 : G/N_1 &\to G/N_2 \\ : gN_1 &\mapsto gN_2\end{aligned}$$

is well-defined. With kernel

$$\pi_{N_1}(N_2) = N_2/N_1.$$

Therefore the First Isomorphism Theorem gives the result. ∎

[V] Denoted simply as π

11.3 Another Important Family of Groups: Projective Linear Groups

For $G \subseteq \mathbf{M}_n(R)$, Recall the definition of $\mathcal{S}G$ from Definition 3.237 back on page 211. Is is the set of scalar matrices which live in G. That is $\mathcal{S}G \stackrel{\text{def}}{=} \mathcal{S}(n,R) \cap G$. Note, for many important examples of matrix groups G, we have that $\mathcal{S}G$ coincides with the center of G. This is the case whenever G is any of the matrix groups

$$\mathbf{GL}_n(R) \quad \mathbf{SL}_n(R) \quad \mathbf{U}_n(\mathbb{C}) \quad \mathbf{O}_n(\mathbb{R}) \quad \mathbf{SO}_n(\mathbb{R}) \quad \mathbf{Sp}_n(R)^{\text{I}}$$

and many others.

Here's an easy way to tell that the center of a matrix group consists of nothing but the scalar matrices in it.

SOLUTION

Exercise 11.11. Let G be a matrix group, i.e. $G \leqslant \mathbf{GL}_n(\mathbb{F})$ (with $\mathbb{F}$ a field). Suppose G contains a diagonal matrix $\mathbf{D}(\mathbf{d})$ with $\mathbf{d}$ consisting of distinct non-zero elements of $\mathbb{F}$. Suppose further that in the action of $G \circlearrowright \mathbb{F}^n$ all standard basis vectors $\langle \mathbf{e}_1 \rangle, \langle \mathbf{e}_2 \rangle \ldots \langle \mathbf{e}_n \rangle$ are contained in the same orbit. (This means for any $i \in \overline{n}$ there exists a $\mathbf{A} \in G$ so that $\mathbf{A}(\mathbf{e}_1) = r\mathbf{e}_i$ for some $r \in \mathbb{F}$.) Prove that

$$\mathcal{C}(G) = \mathcal{S}G.^{\text{II}}$$

Hint! By Exercise 6.27 back on page 337 we know that the center of G is contained inside of the diagonal matrices (since the centralizer of $\mathbf{D}(\mathbf{d})$ has this property. Then for a matrix $\mathbf{A}$ which takes $\mathbf{e}_i$ to a vector proportional to $\mathbf{e}_j$ show we have that

$$(\mathbf{A} \cdot \mathbf{D}(\mathbf{c}) \cdot \mathbf{A}^{-1})(\mathbf{e}_j) = c_i \mathbf{e}_j$$

while

$$\mathbf{D}(\mathbf{c})(\mathbf{e}_j) = c_j \mathbf{e}_j.$$

Hence if $\mathbf{A}$ and $\mathbf{D}(\mathbf{c})$ were to commute we'd have to have $c_i = c_j$.

Example 11.28. This exercise applies often (you might have to use it in pieces if the field $\mathbb{F}$ does not contain enough elements). For instance for $\mathbf{GL}_n(\mathbb{F})$ we can find a diagonal matrix with all distinct entries (again, I'm fudging a bit here, I'll leave it to you to "unfudge") and then use permutation matrices to meet the second part of the hypothesis. Indeed, just the permutation matrix corresponding to the permutation $(1\ 2 \cdots n)$ alone would take care of this in one swoop! For $\mathbf{SL}_n(\mathbb{F})$ you'd need to use even permutation matrices.

In any case, since the matrix $r\mathbf{I}$ represents represents scalar multiplication by r, and since linear maps by their very definition commute with scalar multiplication, we know $\mathrm{D}G \leqslant \mathcal{C}(G)$. In particular, like all subgroups of the center of a group, $\mathcal{S}G \trianglelefteq G$ (Exercise **??** back on page **??**). Therefore, the set of cosets

$$\mathrm{P}G \stackrel{\text{def}}{=} G/\mathcal{S}G$$

forms a group known as *projective* G. For us, the most important special cases are

$$\mathbf{PGL}_n(R) \stackrel{\text{def}}{=} \mathbf{GL}_n(R)/\mathcal{S}\mathbf{GL}_n(R) \quad \text{and} \quad \mathbf{PSL}_n(R) \stackrel{\text{def}}{=} \mathbf{SL}_n(R)/\mathcal{S}\mathbf{SL}_n(R)^{\text{III}}$$

[I] The group $\mathbf{Sp}_n(R)$ is only defined when n is even.
[II] That is show that the only matrices in G which commute with everything in G are scalar matrices.

which are the *projective general linear group* (of degree n and coefficients in R) and the *projective special linear groups* (of degree n and coefficients in R). For starters, it is a remarkable fact that we have that for p-prime the groups $\mathbf{PSL}_n\left(\mathbb{Z}_p\right)$ are simple except in the cases that $n = 2, p = 2$ and $n = 2, p = 3$. This is even true if we replace $\mathbb{Z}_p$ with any finite field (see the appendix on ring theory to see what this means, and how to create such things.) These give us the single largest class of finite simple groups.

To make the notation a bit easier, we use the following convention.

CONVENTION 11.29. We will think of elements of $\mathbf{PGL}_n(R)$ as matrices.

EXERCISE 11.12. Show that for all positive integers n

$$\mathbf{GL}_n(\mathbb{Z}_2) = \mathbf{SL}_n(\mathbb{Z}_2) = \mathbf{PSL}_n(\mathbb{Z}_2).$$

SOLUTION

EXERCISE 11.13. In Exercise 5.3 back on 240 we already defined $\mathbf{PGL}_2(R)$ for several rings R as the set of formal, non-constant rational functions of the form

$$\frac{az+b}{cz+d}.$$

SOLUTION

Show that this new definition of $\mathbf{PGL}_2(R)$ and the old definition for $\mathbf{PGL}_2(R)$ are isomorphic.
Hint! Let's denote the old definition of $\mathbf{PGL}_2(R)$ as $\overline{\mathbf{PGL}_2(R)}$. Define a function

$$\begin{aligned}\varphi\colon \mathbf{GL}_2(R) &\to \overline{\mathbf{PGL}_2(R)}\\ \colon \begin{pmatrix} a & b\\ c & d\end{pmatrix} &\mapsto \frac{az+b}{cz+d}.\end{aligned}$$

Use Exercise 8.18, and then use the first isomorphism theorem.

EXERCISE 11.14.

SOLUTION

(1) Prove that there are $\mathbf{O}_n(\mathcal{U}(\mathbb{Z}_p)) = \gcd(n, p-1)$ elements of order n in $\mathcal{U}(\mathbb{Z}_p)$.
Hint! Use two theorems and prove this in 1 line.

(2) Prove that $\mathrm{D}\mathbf{SL}_n\left(\mathbb{Z}_p\right)$ also contains $\gcd(n, p-1)$ elements.

(3) Prove that

$$\left|\mathbf{PSL}_n\left(\mathbb{Z}_p\right)\right| = \frac{p^{\binom{n}{2}}}{\gcd(n, p-1)} \cdot \prod_{i=2}^{n}(p^i - 1).$$

[III] Every time I order a pumpkin spice late from starbucks, they write PSL on my cup. Who knew they were math nerds?

(4) Give a more explicit formula for the order of $\mathbf{PSL}_2\left(\mathbb{Z}_p\right)$.

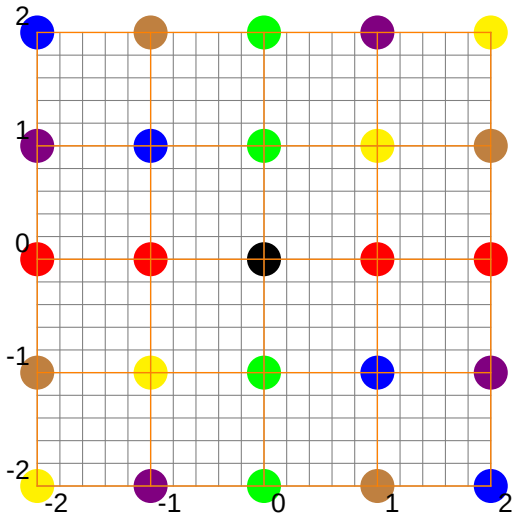
(5) What is the order of $\mathbf{PSL}_3\left(\mathbb{Z}_7\right)$?

Recall (first discussed back on page 50) that two points $\mathbf{v}$ and $\mathbf{w}$ in $V = \mathbb{Z}_p^n \setminus \{\mathbf{0}\}$ are said to live in the same *line* if there exists a non-zero scalar, $c \in \mathbb{Z}_p$, with $c\mathbf{v} = \mathbf{w}$. The defining equation gives rise to an equivalence relation defined on V whose classes are the set of lines. In particular, every point in V (i.e. every non-zero vector) lies in a unique line. If $\mathbf{v} \in V$ we denote the line containing V as $\langle\, \mathbf{v} \,\rangle$. For instance, in the case that $p = 5$ and $n = 2$ the lines are

$$\left\langle \begin{bmatrix}1\\0\end{bmatrix} \right\rangle := \left\{ \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}2\\0\end{bmatrix}, \begin{bmatrix}3\\0\end{bmatrix}, \begin{bmatrix}4\\0\end{bmatrix} \right\} \qquad \left\langle \begin{bmatrix}1\\1\end{bmatrix} \right\rangle := \left\{ \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}2\\2\end{bmatrix}, \begin{bmatrix}3\\3\end{bmatrix}, \begin{bmatrix}4\\4\end{bmatrix} \right\}$$

$$\left\langle \begin{bmatrix}1\\2\end{bmatrix} \right\rangle := \left\{ \begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\4\end{bmatrix}, \begin{bmatrix}3\\1\end{bmatrix}, \begin{bmatrix}4\\3\end{bmatrix} \right\} \qquad \left\langle \begin{bmatrix}1\\3\end{bmatrix} \right\rangle := \left\{ \begin{bmatrix}1\\3\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix}, \begin{bmatrix}4\\2\end{bmatrix} \right\}$$

$$\left\langle \begin{bmatrix}1\\4\end{bmatrix} \right\rangle := \left\{ \begin{bmatrix}1\\4\end{bmatrix}, \begin{bmatrix}2\\3\end{bmatrix}, \begin{bmatrix}3\\2\end{bmatrix}, \begin{bmatrix}4\\1\end{bmatrix} \right\} \qquad \left\langle \begin{bmatrix}0\\1\end{bmatrix} \right\rangle := \left\{ \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}0\\2\end{bmatrix}, \begin{bmatrix}0\\3\end{bmatrix}, \begin{bmatrix}0\\4\end{bmatrix} \right\}.$$



A more group theoretical way of talking about lines is by considering the set of subgroups of $\mathbb{Z}_p^n$ of order p. Note that since every element of $\mathbb{Z}_p^n$ has order p every element of this group is in such a subgroup. Moreover no distinct pair of these subgroups can meet at any point other than the identity. This is because if two subgroups of a group of order p were to intersect, then since any element of a group of order p generates it, two subgroups of a group of order p which have a non-trivial intersection must contain a generator in common. Thus they must contain everything in common. Thus they really are the same subgroup.

EXERCISE 11.15. Show that there are a total of

$$1 + p + p^2 + \cdots + p^{n-1}$$

lines in $\mathbb{Z}_p^n$.

Hint! This is pretty easy. Indeed anyway you look at it, it should be clear that each line in $\mathbb{Z}_p^n$ contains exactly $p - 1$ points (not equal to $\mathbf{0}$).

SOLUTION

EXERCISE 11.16. Show that $\mathbf{PSL}_2\left(\mathbb{Z}_p\right)$ acts

(1) Faithfully

(2) Transitively

on the set of $p + 1$ lines in $\mathbb{Z}_p \times \mathbb{Z}_p = \mathbb{Z}_p^2$.

SOLUTION

EXAMPLE /THEOREM 11.30. $\mathcal{A}_4 \cong \mathbf{PSL}_2(\mathbb{Z}_3)$

PROOF. We have previously seen in Example 10.76 that that $B\left(\blacktriangleright^3\right) \cong \mathbf{SL}_2(\mathbb{Z}_3)$. But since $B\left(\blacktriangleright^3\right)/\{\pm 1\} \cong S\left(\blacktriangleright^3\right) \cong \mathcal{A}_4$ we must also have that $\mathbf{SL}_2(\mathbb{Z}_3)$ mod the corresponding subgroup of matrices is also isomorphic to $\mathcal{A}_4$. But this corresponding subgroup is obviously $\pm\mathbf{I}_2$ and so this quotient is, by its very definition, $\mathbf{PSL}_2(\mathbb{Z}_3)$. ■

EXAMPLE /THEOREM 11.31. $\mathcal{A}_5 \cong \mathbf{PSL}_2(\mathbb{Z}_5)$.

PROOF. $\mathcal{A}_5$ is generated by the 5–cycles

$$\sigma := (1\,2\,3\,4\,5) \qquad \tau := (1\,3\,4\,2\,5).$$

This can be seen in many ways, the easiest is by Theorem 10.136 back on page 590 which says this outright, since σ and τ do not have any powers in common (except for the ones that each yield the identity).

The product of these 5–cycles has order 5. Under the action from Example 10.78 back on page 541 what do these 5–cycles correspond to? Well, the first one, $\sigma = (1\,2\,3\,4\,5)$ is done out in that example. There we said it went to

$$(H_1)(H_2\,H_6\,H_5\,H_4\,H_3)$$

We also showed that

$$(4\,5) \mapsto (H_1\,H_4)(H_5\,H_6)(H_2\,H_3)$$

To figure out where $\tau = (1\,3\,4\,2\,5)$ ends up, we note that

$$\begin{aligned}
(1\,2\,3\,4\,5)^2 = (1\,3\,5\,2\,4) &\Rightarrow (4\,5)(1\,2\,3\,4\,5)^2(4\,5)^{-1} = (1\,3\,4\,2\,5) = \tau \\
&\Rightarrow \tau \mapsto (H_1\,H_4)(H_5\,H_6)(H_2\,H_3)\Big[(H_1)(H_2\,H_6\,H_5\,H_4\,H_3)^2\Big](H_1\,H_4)(H_5\,H_6) \\
&= (H_1\,H_4)(H_5\,H_6)(H_2\,H_3)\Big[(H_1)(H_2\,H_5\,H_3\,H_6\,H_4)\Big](H_1\,H_4)(H_5\,H_6) \\
&= (H_4)(H_3\,H_6\,H_2\,H_5\,H_1) \\
&= (H_4)(H_1\,H_3\,H_6\,H_2\,H_5).
\end{aligned}$$

In summary

$$(1\,2\,3\,4\,5) \mapsto (H_1)(H_2\,H_6\,H_5\,H_4\,H_3) \quad \text{and} \quad (1\,3\,4\,2\,5) \mapsto (H_4)(H_1\,H_3\,H_6\,H_2\,H_5).$$

Meanwhile the generators $\mathfrak{R}$ and $\mathfrak{L}$ of $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$, given as usual as

$$\mathfrak{L} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathfrak{R} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

act on the six lines in $\mathbb{Z}_5^2$ as follows:

$$\begin{aligned} \mathfrak{L} \bullet \langle (1,0) \rangle &= \langle (1,0) \rangle \\ \mathfrak{L} \bullet \langle (0,1) \rangle &= \langle (1,1) \rangle \\ \mathfrak{L} \bullet \langle (1,3) \rangle &= \langle (4,3) \rangle = \langle (1,2) \rangle \end{aligned} \qquad \begin{aligned} \mathfrak{L} \bullet \langle (1,1) \rangle &= \langle (2,1) \rangle = \langle (1,3) \rangle \\ \mathfrak{L} \bullet \langle (1,2) \rangle &= \langle (3,2) \rangle = \langle (1,4) \rangle \end{aligned}$$

while

$$\begin{aligned} \mathfrak{R} \bullet \langle (1,0) \rangle &= \langle (1,1) \rangle \\ \mathfrak{R} \bullet \langle (1,2) \rangle &= \langle (1,3) \rangle \\ \mathfrak{L} \bullet \langle (0,1) \rangle &= \langle (0,1) \rangle . \end{aligned} \qquad \begin{aligned} \mathfrak{R} \bullet \langle (1,1) \rangle &= \langle (1,2) \rangle \\ \mathfrak{R} \bullet \langle (1,3) \rangle &= \langle (1,4) \rangle \end{aligned}$$

We use the obvious dictionary in the first case, that is

$$\begin{aligned} 1 &= H_1 \\ 4 &= H_4 \end{aligned} \qquad \begin{aligned} 2 &= H_2 \\ 5 &= H_5 \end{aligned} \qquad \begin{aligned} 3 &= H_3 \\ 6 &= H_6. \end{aligned}$$

For things to work out, i.e. for

$$\sigma \mapsto \mathfrak{R} \quad \text{and} \quad \tau \mapsto \mathfrak{L}$$

to be an isomorphism, via all this we'd need to define a dictionary on the set of lines via

$$\begin{aligned} 1 &= \langle (1,0) \rangle \\ 4 &= \langle (0,1) \rangle \end{aligned} \qquad \begin{aligned} 2 &= \langle (1,3) \rangle \\ 5 &= \langle (1,4) \rangle \end{aligned} \qquad \begin{aligned} 3 &= \langle (1,1) \rangle \\ 6 &= \langle (1,2) \rangle . \end{aligned}$$

Indeed, a direct check now shows that under these dictionaries both $\mathfrak{L}$ and σ act via $(1)(2\,6\,5\,4\,3)$ and both $\mathfrak{R}$ and τ act via $(4)(1\,3\,6\,2\,5)$. Therefore we have both

$$\mathcal{A}_5 \xrightarrow{\cong} \langle (2\,6\,5\,4\,3), (1\,3\,6\,2\,4) \rangle \xleftarrow{\cong} \mathbf{PSL}_2(\mathbb{Z}_5) \tag{2}$$

and hence the groups $\mathcal{A}_5$ and $\mathbf{PSL}_2(\mathbb{Z}_5)$ are isomorphic to each other. ■

Remark 11.32. Technically, I didn't prove that the first map in (2) was a bijection (it is a surjection since the middle group is by definition, generated by the image of elements in $\mathcal{A}_5$). In other words, I didn't prove that the action of $\mathcal{A}_5$ on the 6 things was faithful. However, we know the action of $\mathbf{PSL}_2(\mathbb{Z}_5)$ on the 6 things is faithful (only a multiple of the identity matrix can fix all lines), so we have indeed proven the second isomorphism in (2). Therefore the middle group, being isomorphic to a group of order 60, does indeed have order 60. Therefore, the first map in that equation is a surjection from a set of order 60 to another set of order 60, therefore, by the pigeon hole principle, it must also be injective, so a bijection. Therefore, since it's a permutation representation it's also a homomorphism, hence it's a bijective homomorphism, i.e. an isomorphism.

EXAMPLE /THEOREM 11.33. For any prime p, and integer n, the groups $\mathbf{PGL}_n\left(\mathbb{Z}_p\right)$ and $\mathbf{PSL}_n\left(\mathbb{Z}_p\right)$ are centerless.

PROOF. The n by n matrix

$$\mathbf{A}_n = \mathbf{A} := \begin{pmatrix} 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

[IV] can easily be checked to only fix the line

$$\langle\, \mathbf{e}_1 \,\rangle = \left\langle \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\rangle.$$

Meanwhile, it's easy to see that $\mathbf{PSL}_n\left(\mathbb{Z}_p\right)$ (and hence also the larger group $\mathbf{PGL}_n\left(\mathbb{Z}_p\right)$) acts transitively on the set of lines in $\mathbb{Z}_p^n$, $\left(\mathbb{Z}_p^n \setminus \{(0,\dots 0)\}\right)_{/\asymp}$. Therefore for the faithful actions of $\mathbf{PSL}_n\left(\mathbb{Z}_p\right)$ and $\mathbf{PGL}_n\left(\mathbb{Z}_p\right)$ on the set of lines, there exists a matrix which fixes that line and only that line. Therefore, by Theorem 7.29, the groups $\mathbf{PSL}_n\left(\mathbb{Z}_p\right)$ and $\mathbf{PGL}_n\left(\mathbb{Z}_p\right)$ have been proven to be centerless, as promised. ■

We prove something that doesn't fit with this chapter, but will be useful for you in the exercise below.

LEMMA 11.34. *Let* $\mathbf{A} \in \mathbf{SL}_2\left(\mathbb{Z}_p\right)$ *satisfies* $\mathbf{A}^p = \pm\mathbf{I}_2$, *then* $\mathbf{A}$ *has trace* ± 2.

PROOF. Taking inverses of both sides of the equation $\mathbf{A}^p = \pm\mathbf{I}_2$ implies that we also have that $\left(\mathbf{A}^{-1}\right)^p = \pm\mathbf{I}_2$ as well. Now any element of any group group commutes with its inverse. Therefore we may apply a Freshmen's Dream from the appendix, Theorem K.3 coming up on page 878 to imply that

$$(\mathbf{A} + \mathbf{A}^{-1})^p = \mathbf{A}^p + \mathbf{A}^{-p} = \pm 2\mathbf{I}_2.$$

But since $\det \mathbf{A} = 1$ we know that $\mathbf{A} + \mathbf{A}^{-1} = \operatorname{tr}(\mathbf{A})\,\mathbf{I}_2$. Plugging this in shows us that

$$\operatorname{tr}(\mathbf{A})^p\,\mathbf{I}_2 = \pm 2\mathbf{I}_2.$$

But by Fermat's little theorem, $\operatorname{tr}(\mathbf{A})^p = \operatorname{tr}(\mathbf{A})$. Therefore we have prove that $\operatorname{tr}(\mathbf{A}) = \pm 2$ as desired. ■

[IV] So for $n = 3$, $\mathbf{A}_3$ is

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

and for $n = 4$, $\mathbf{A}_4$ is

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

In general, $\mathbf{A}_n$ can be thought of as the identity matrix + the matrix with ones directly above the diagonal, i.e. the linear map $\mathbf{e}_i \mapsto \mathbf{e}_{i-1}$ for $i > 1$ and $\mathbf{e}_1 \mapsto \mathbf{0}$.

Actually, if one replaces the $\mathbf{A}^{-1}$ with the adjoint of $\mathbf{A}$, $\widetilde{\mathbf{A}}$, in the proof above, one can prove a stronger result for $\mathbf{A}$ of any determinant, namely that $\mathbf{A}^p$ is a diagonal matrix iff

$$\mathrm{tr}^2(\mathbf{A}) = 4 \cdot \det(\mathbf{A}).$$

We generalize this greatly in Appendix G which starts on 837.

Recall that ${}^{\pm}\mathbf{SL}_n(R)$ is the group of n by n matrices with coefficients in $\mathbb{Z}_n$ and with determinant equal to $[-1]_n$ or $[+1]_n$. To get our thoughts organized for the next exercise, note that $\mathcal{S}^{\pm}\mathbf{SL}_2(\mathbb{Z}_n)$ is generated by

$$\begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$$

if $\mathbb{Z}_n$ contains an i (i.e. an element $i \in \mathbb{Z}_n$ with $i^2 \equiv_n -1$) and by

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

otherwise.

SOLUTION

EXERCISE **Hard!** **11.17.** Let $\mathbf{\Phi}$ be the *golden matrix*

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Note that $\mathbf{\Phi}$ happens to have determinant -1. Unlike our usual convention, we think of the matrix $\mathbf{\Phi}$ as having integer coefficients (so that $\mathbf{\Phi} \in^{\pm} \mathbf{SL}_2(\mathbb{Z})$ (i.e. $\mathbf{GL}_2(\mathbb{Z})$) and we explicitly write $[\mathbf{\Phi}]_n$ when we wish to think of the coefficients of the matrix as living in $\mathbb{Z}_n$.

(1) Prove that $[\mathbf{\Phi}]_m^n$ is the identity in $\mathbf{P}^{\pm}\mathbf{SL}_2(\mathbb{Z}_m)$ (i.e. it is a scalar matrix) iff $m \mid F_n$, where as usual, F_n is the n^{th} term in the Fibonacci sequence defined on 101. [V]
Hint! Reprove or look up: $\mathbf{\Phi}^n = F_n \cdot \mathbf{\Phi} + F_{n-1} \cdot \mathbf{I}_2$.

(2) Prove that the order of $[\mathbf{\Phi}]_m$ in $\mathbf{P}^{\pm}\mathbf{SL}_2(\mathbb{Z}_m)$ is the least positive integer n so that $m \mid F_n$.

(3) Prove that every integer divides some term in the Fibonacci sequence.
Hint! Previous part and Corollary 5.52 applied to $\mathbf{P}^{\pm}\mathbf{SL}_2(\mathbb{Z}_m)$.

(4) Use the previous parts to prove that

$$\gcd(F_n, F_m) = F_{\gcd(n,m)}.$$ [VI]

Hint! This is not too too hard, but it takes **a lot** of focus! Try to take this one step at a time. Perhaps checking out Appendix K, which starts on page 877 will help your confidence.

(5) Prove that for a prime $p \neq 5$, p does not divide F_p.
Hint! As usual in number theory, dispense with the odd (odd meaning strange. hahaha, get it?) case of $p = 2$ first. Now assume that p is an odd prime. If $p \mid F_p$, then $[\mathbf{\Phi}]_p^{2p}$ is either $\mathbf{I}_2$ or $-\mathbf{I}_2$ (in $\mathbf{P}^{\pm}\mathbf{SL}_2\left(\mathbb{Z}_p\right)$) (actually a bit more thought would show that it has to be the latter). Then use the lemma to prove that $\mathbf{\Phi}^2$ must have trace 2. Lastly, by explicitly using the matrix $\mathbf{\Phi}$, show that this happens only for $p = 5$.

(6) Let n be an odd number and suppose that p is a prime other than 2 which divides F_n. Prove that $p \equiv_4 1$.
Hint! Prove that if $[\mathbf{\Phi}]_p^n$ is a scalar matrix, then $\mathbb{Z}_p$ has an i. Use Example/Theorem 5.86 back on page 293.

11.4 Subgroup Correspondence

Definition 11.35. Suppose $H \leqslant G$. Define

$$\mathcal{G}(G,H) \stackrel{\text{def}}{=} \{K : H \leqslant K \leqslant G\}$$

and

$$\mathcal{G}(G) \stackrel{\text{def}}{=} \mathcal{G}(G,\{e_G\}) = \text{Set of all subgroups of } G.$$

Theorem 11.36 (Subgroup Correspondence)**.** *Suppose that $N \trianglelefteq G$. Then*

$$H \mapsto {}^{H}\!/_{N}$$

defines a bijection

$$\Psi\colon \mathcal{G}(G,N) \to \mathcal{G}(G/N)$$

which preserves

(1) *Order.* (*That is $H_1 \leqslant H_2$ for subgroups of G containing N iff $H_1/N \leqslant H_2/N$.*)

(2) *Index.* (*That is if $H \in \mathcal{G}(G,N)$ has index n in G iff H/N has index n in G/N.*)

(3) *Normal Subgroups.* (*That is if $N \leqslant \widehat{N}$, then $\widehat{N} \trianglelefteq G$ iff $\widehat{N}/N \trianglelefteq G/N$.*)

Proof. I claim that inverse of the map Ψ is the map

$$\Phi\colon \overline{H} \leqslant G/N \mapsto \pi_N^{-1}(\overline{H}) = \{g \in G : gN \in \overline{H}\}.$$

Since $N = e_{G/N} \in \overline{H}$, we have that

$$N = \pi_N^{-1}(N) \leqslant \pi_N^{-1}(\overline{H})$$

we have that π_N really has codomain $\mathcal{G}(G,N)$. Next note that like for all surjective functions,

$$\pi_N(\pi_N^{-1}(\overline{S})) = \overline{S}$$

for all subsets $\overline{S} \subseteq G/N$. This means that $\Psi \circ \Phi = \mathbf{1}_{\mathcal{G}(G/N)}$. On the other hand, if $H \leqslant G$, then like for any homomorphism, $\pi_N^{-1}(\pi_N(H)) = HN$. However if we assume moreover that $N \leqslant H$, then $HN = H$. Thus, restricted to $\mathcal{G}(G,N)$, we have that $\Phi \circ \Psi = \mathbf{1}_{\mathcal{G}(G/N)}$ as well. Therefore we have proven that Ψ is a bijection and that Φ is its explicit inverse. Let's verify that Ψ preserves each thing indicated.

(1) For any bijection $\sigma\colon X \to Y$ between any two sets, $\sigma(X_1) \subseteq \sigma(X_2)$ holds iff $X_1 \subseteq X_2$. Therefore the fact that Ψ is a bijection automatically means that it preserves $\subseteq$, which is all the order is on $\mathcal{G}(G,N)$.

[V] Compare with the more complicated result when we look in $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ instead of $\mathbf{P}^{\pm}\mathbf{SL}_2\left(\mathbb{Z}_p\right)$, c.f. Exercise 5.30

[VI] For instance, this says that for p a prime, F_p is not divisible by F_n for any $n < p$ and for $n > 4$ not prime, F_n is not prime.

(2) For $H \in \mathcal{G}(G,N)$, we know that $\Psi(H) \cong H/N$. Therefore

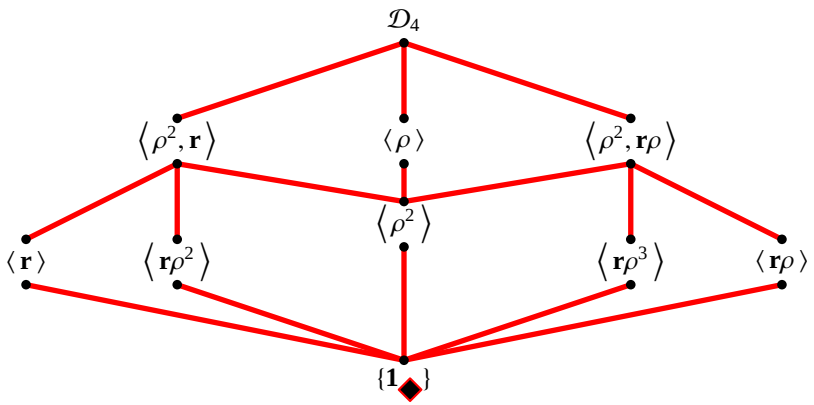
$$\begin{aligned}
[G:N] &= [G:H]\cdot[H:N] \\
&= [G:H]\cdot|\Psi(H)| \\
&= [G:H]\cdot|\Psi(H)|\cdot\frac{|G/N|}{|G/N|} \\
&\Rightarrow |G/N| = [G:H]\cdot|\Psi(H)| \\
&\Rightarrow [G/N:\Psi(H)] = |G/N|/|\Psi(H)| = [G:H].
\end{aligned}$$

(3) Another direct check {Clean up}!

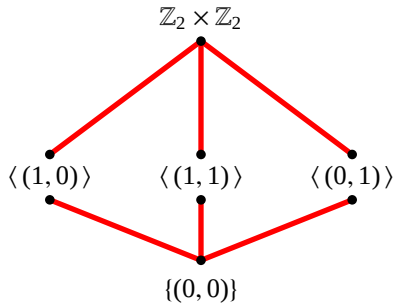
■

This means if we draw the subgroup lattice of G in say red and color all subgroups of (and arrows between such subgroups) G which contain N in blue, then the red sub poset would be exactly the same as the subgroup lattice of G/N. For example:

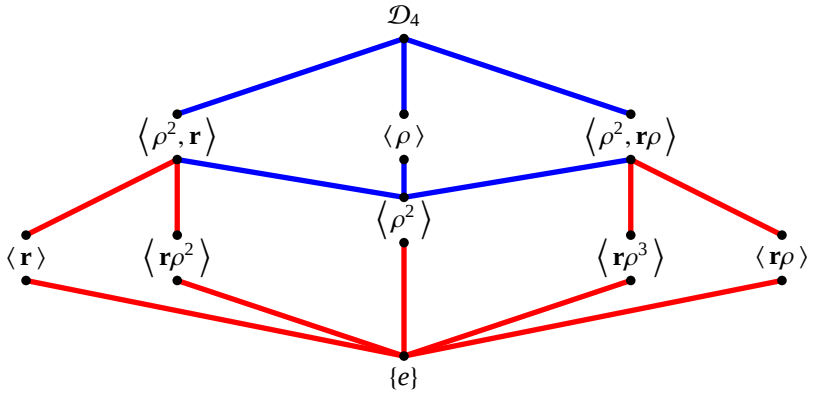
EXAMPLE 11.37. Let $G = \mathcal{D}_4$ and $N = \left\langle \rho^2 \right\rangle$. The subgroup lattice of $\mathcal{D}_4$ is depicted as follows:



We have seen a few times that $\mathcal{D}_4/N$ is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2 = \mathbb{Z}_2^2$ which has a subgroup lattice



The content of the theorem is that when we highlight the part of the first diagram "above" $\langle \rho^2 \rangle$



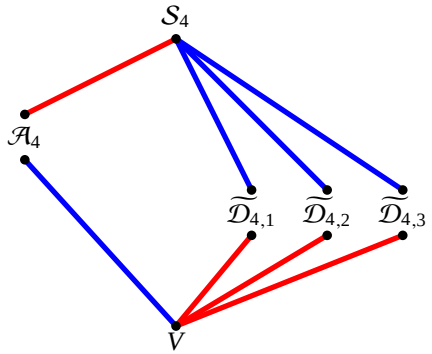
the highlighted sublattice is "the same as" the lattice diagram for $\mathbb{Z}_2^2$.

EXAMPLE 11.38. Consider G to be $\mathcal{S}_4$ and N to be the Klein four group, $V \stackrel{\text{def}}{=} \langle (1\,2)(3\,4), (1\,3)(2\,4) \rangle \trianglelefteq \mathcal{S}_4$. We verify that $\mathcal{G}(\mathcal{S}_4, V)$ and its quotient, which is isomorphic to $\mathcal{S}_3$, have the same lattice diagrams.

For starters, there are a total of 6 subgroups of $\mathcal{S}_4$ which contain V, namely V itself, the three copies of the symmetries of the square, $\mathcal{D}_4$, in $\mathcal{S}_4$, the alternating group, $\mathcal{A}_4$, and $\mathcal{S}_4$ itself.[I] In more detail:

$$\underset{V}{\widetilde{\mathcal{D}}_{4,1} = \langle (1\,2\,3\,4), (2\,4) \rangle} \qquad \underset{\mathcal{A}_4}{\widetilde{\mathcal{D}}_{4,2} = \langle (1\,3\,4\,2), (1\,4) \rangle} \qquad \underset{\mathcal{S}_4.}{\widetilde{\mathcal{D}}_{4,3} = \langle (1\,3\,2\,4), (3\,4) \rangle}$$

They make up the lattice

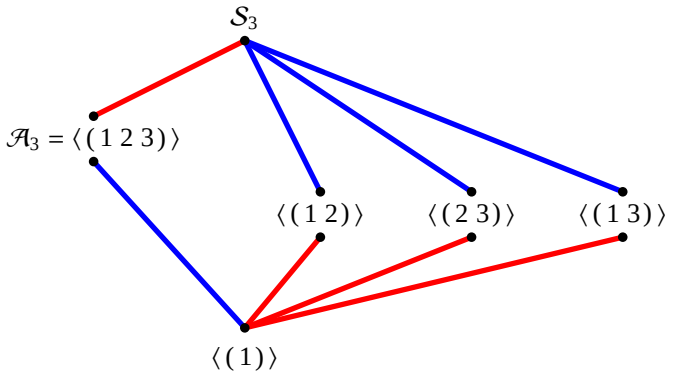


[I] Perhaps you don't believe this. In which case you can think of this example as proving that fact.

(Here the blue lines indicate index 3 and the red lines indicate index 2) Meanwhile, $\mathcal{S}_4/V \cong \mathcal{S}_3$ via the isomorphism

$$\begin{aligned}(1\,2)V &\mapsto (1\,2)\\ (2\,3)V &\mapsto (2\,3)\\ (3\,4)V &\mapsto (1\,2)\,.\end{aligned}$$

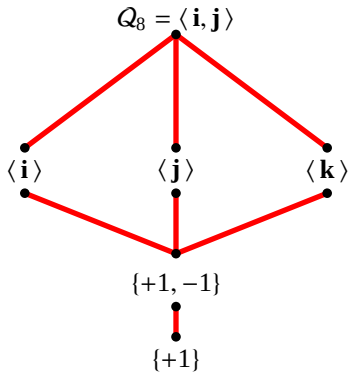
Now $\mathcal{S}_3$ has subgroup a lattice



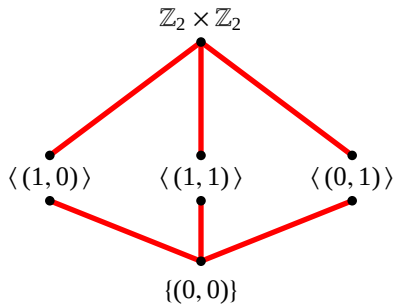
.

Of course these two diagrams match perfectly, which is in accordance with the theorem. In fact this proves, without a shadow of a doubt, that the the 6 subgroups listed above are the only subgroups of $\mathcal{S}_4$ which contain V, a fact that could be verified in other ways, but with little fun.

EXAMPLE 11.39. In Example 10.74 back on page 535 we showed the subgroup lattice of Q_8 to be



. Notice that this is the same lattice as the subgroup lattice for $\mathbb{Z}_2^2$.



except the lattice for Q_8 has a tail stuck onto the end given by the trivial group. We now know why this is. It has to do with the following general theorem which says that the same correspondence holds between the double dihedral groups and the dihedral groups of degree a power of 2. (Recall that $Q_8 = B\mathcal{D}_2$ and $\mathbb{Z}_2^2 = \mathcal{D}_2$.)

EXAMPLE /THEOREM 11.40. For every integer $i > 0$, the subgroup lattice of the non-trivial subgroups of $B\mathcal{D}_{2^i}$ is the same as the lattice of all subgroups of $\mathcal{D}_{2^i}$.

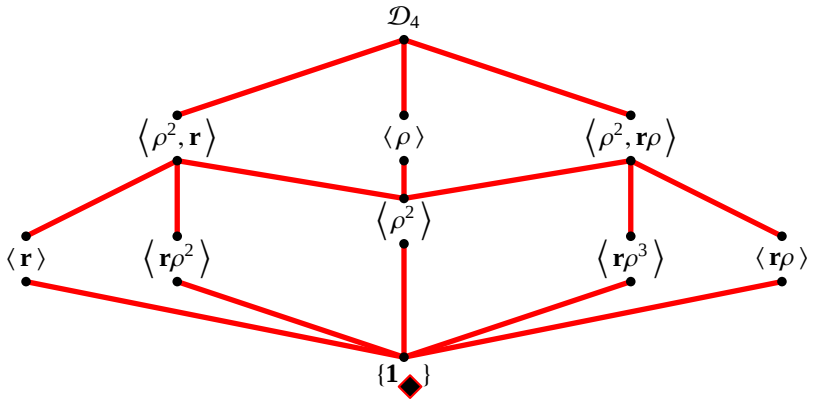
PROOF. We have an epimorphism

$$\mathbf{R}\colon B\mathcal{D}_{2^i} \to \mathcal{D}_{2^i}{}^{\text{II}}$$

whose kernel consists only of $\{\pm 1\} = \langle -1 \rangle$. By the first isomorphism theorem, we can identify $\mathcal{D}_{2^i}$ with $B\mathcal{D}_{2^i}/\pm 1$. Therefore, by the subgroup correspondence theorem, Theorem 11.36, to prove the current example/theorem, we simply need to prove that every non-trivial subgroup of $B\mathcal{D}_{2^i}$ contains -1.

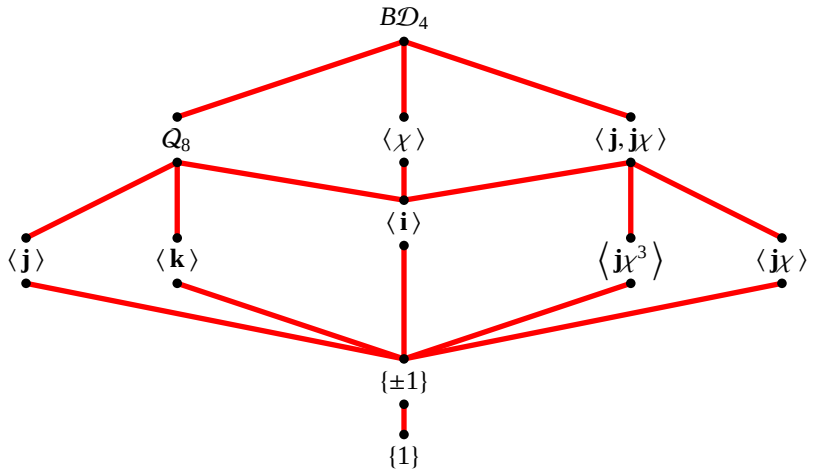
Indeed, since $B\mathcal{D}_{2^i}$ is a 2–group (its order is $4 \cdot 2^i$) and hence so are all of its non-trivial subgroups. As such, each non-trivial subgroup of $B\mathcal{D}_{2^i}$ must contain an element of order 2. But the only quaternion of order 2 is -1. Hence all non-trivial subgroup of $B\mathcal{D}_{2^i}$ must contain -1 which is what we wanted to prove. ■

To see another example of this, let's check out the subgroup lattice of $B\mathcal{D}_4$, which is the double of $\mathcal{D}_4$, the symmetries of the square. We just showed that the subgroup lattice of $\mathcal{D}_4$ was



Therefore the lattice for $B\mathcal{D}_4$ is

[II] This homomorphism takes $\chi \mapsto \rho$ and $j \mapsto \mathbf{r}$.



11.4.1 Subgroups of Cyclic Groups

To give us more suggestive notation, let's denote the cyclic group generated by an element $[\,d\,]_n \in \mathbb{Z}_n$ by $d\mathbb{Z}_n$. That is

$$d\mathbb{Z}_n := \{d\,[\,m\,]_n : [\,m\,]_n \in \mathbb{Z}_n\} = \{[\,d\,]_n\,[\,m\,]_n : [\,m\,]_n \in \mathbb{Z}_n\} = \langle\, [\,d\,]_n \,\rangle\,.$$

Note since homomorphisms take the take cyclic group generated by an element of the domain of a homomorphism to the cyclic group generated by the image of that element, we have that for any integer $m \in \mathbb{Z}$

$$\pi_n(m\mathbb{Z}) = m\mathbb{Z}_n.$$

Also note that

$$\begin{aligned}
m\mathbb{Z}_n &= \pi(\pi^{-1}(m\mathbb{Z}_n)) && (\pi \text{ is surjective}) \\
&= \pi(\pi^{-1}(\pi(m\mathbb{Z}))) && (\forall m \in \mathbb{Z}, \pi(m\mathbb{Z}) = m\mathbb{Z}_n) \\
&= \pi(m\mathbb{Z} + n\mathbb{Z}) && (\text{Example/Theorem 8.17 written additively}) \\
&= \pi(\gcd(m,n)\mathbb{Z}) && (\text{Theorem 3.31}) \\
&= \gcd(m,n)\mathbb{Z}_n. && (\forall m \in \mathbb{Z}, \pi(m\mathbb{Z}) = m\mathbb{Z}_n)
\end{aligned}$$

By the way, this gives us a different proof that m and $\gcd(m,n)$ have the same order, which is clearly $\frac{n}{\gcd(n,m)}$.

The following theorem refines Theorem 6.22 back on page 319.

Theorem 11.41. *Every subgroup of $\mathbb{Z}_n$ is equal to $d\mathbb{Z}_n$ for some $d \mid n$.*

Proof. Since $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$, the theorem says that the subgroups of $\mathbb{Z}_n$ are the same as the subgroups of $\mathbb{Z}$ which contain $n\mathbb{Z}$. But every subgroup of $\mathbb{Z}$ is cyclic, and has the form $d\mathbb{Z}$. Such a subgroup $d\mathbb{Z}$ contains $n\mathbb{Z}$ iff $d \mid n$. Lastly, the correspondence of subgroups of $\mathbb{Z}$ containing $n\mathbb{Z}$ and the subgroups of $\mathbb{Z}_n$ (defined in the section) is defined by taking a subgroup of $\mathbb{Z}$ containing $n\mathbb{Z}$ to its image. But we just showed that these images are $d\mathbb{Z}_n$ where $d \mid n$. ■

COROLLARY 11.42. *Every subgroup of a cyclic group, C, is cylic. Moreover if C is finite, these exists a unique subgroup of order d for every divisor of $|C|$.*

PROOF. Infinite cyclic groups are isomorphic to $\mathbb{Z}$ and finite cyclic groups are isomorphic to $\mathbb{Z}_n$, where n is their order. We have proven these statements for these groups and so we have proven these statements for general cyclic groups. ■

This question doesn't really fit in here, but I couldn't think where else to put it.

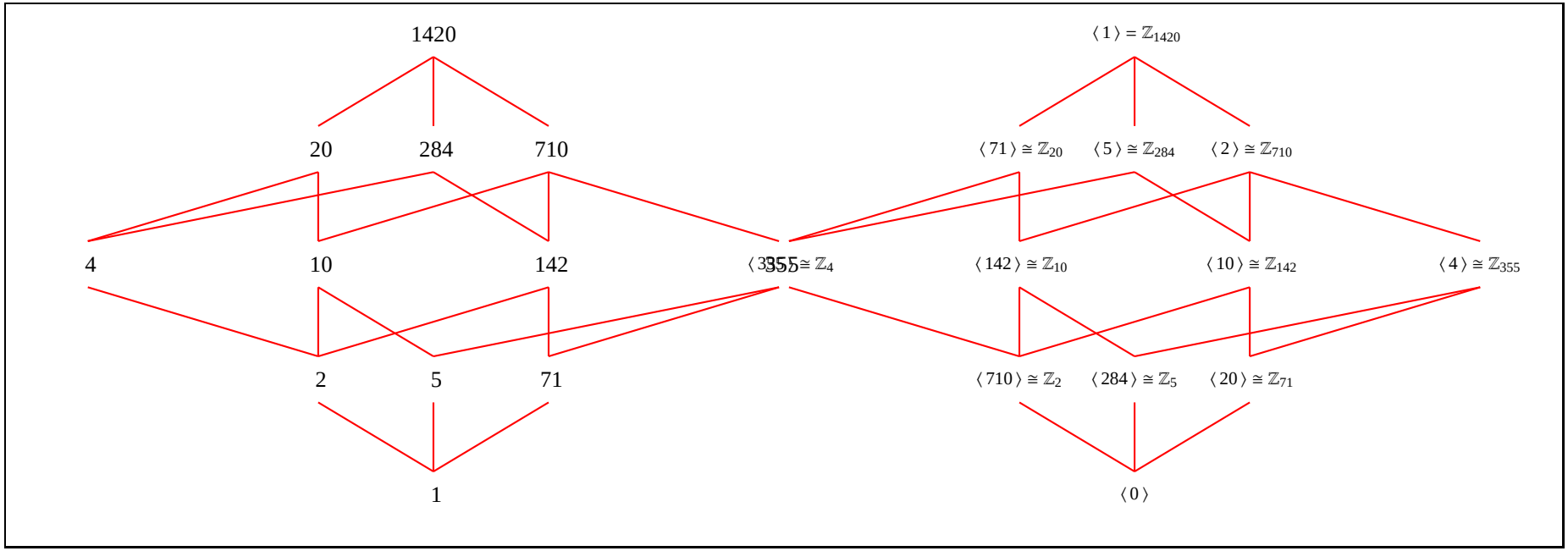
EXERCISE 11.18. That for any divisor $d \mid n$ and $a \in \mathbb{Z}_n$ that $d \cdot a = [0]_n$ iff there exists $m \in \mathbb{Z}_n$ so that $\frac{n}{d}m = a$.

SOLUTION

Since all cyclic groups of order n are isomorphic to $\mathbb{Z}_n$, the previous problem applied to the cyclic groups of order $p-1$, $\mathcal{U}(\mathbb{Z}_p)$, says that:

COROLLARY 11.43. *An element $m \in \mathcal{U}(\mathbb{Z}_p)$ is a perfect square iff $m^{\frac{(p-1)}{2}} = 1$.*

This means that the subgroup lattice of a finite cyclic group of order n is the same as the lattice of divisors of n (see page 51). For instance the lattice of divisors of 1420 is depicted in Figure 11.4.1 and the lattice of subgroups of $\mathbb{Z}_{1420}$ is depicted below.



11.5 Abelianization

The group $\mathcal{D}_4$ contains three, maximal, abelian, proper subgroups; namely $\langle a \rangle$, $\left\langle a^2, b \right\rangle$, and $\left\langle a^2, ab \right\rangle$. The first of which is isomorphic to $\mathbb{Z}_4$ and the second two are isomorphic to $\mathbb{Z}_2^2$, the Klein four group. Given this sad state of things, we see that it doesn't make sense to talk about the largest abelian subgroup inside of a group. We are however, lucky enough to have a largest abelian quotient in a sense discussed in this section.

Notation 11.44. For any two elements a and b of a group G,

$$[a, b] \stackrel{\text{def}}{=} aba^{-1}b^{-1}. \tag{3}$$

We've seen already that if aN and bN commute in G/N (for some $N \trianglelefteq G$) iff $[a, b] \in N$. Therefore, if G/N is abelian, it needs to contain all commutators,

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

(The only problem is that we don't know that the set of all of these commutators forms a subgroup of G.) This motivates the following definition

Definition 11.45. The *derived subgroup* or *commutator subgroup* of a group is the group generated by all commutators. In symbols

$$[G, G] := \langle \{ [a, b] : a, b \in G \} \rangle .$$

Notice, since $[G, G]$ is defined to be the group generated by a subset of G, there is no need to actually prove that $[G, G]$ is a subgroup of G! It's true by definition!

Warning! 11.46. The set of commutators do not in general form a subgroup of G. [I]There is no reason to expect that the set of all commutators to be closed under multiplication. This is why $[G, G]$ was defined to be the group *generated* by the set of commutators and not just the set of commutators themselves.

SubDefinition 11.47. Let $H, K \leqslant G$. Define

$$[H, K] \stackrel{\text{def}}{=} \langle \{ [h, k] : h \in H, k \in K \} \rangle .$$

Again, I should warn you that $[H, K]$ is **not** the set of commutators, but it **is** the group generated by the set of those commutators. This is one of the most common mistakes that students make. As a matter of fact, I think that it's a mistake that nearly every student makes at least once! So you might as well make it and get it over with,

[I]Actually they often do form a subgroup! In fact the order of the group where this fails for the first time is 96. In other words for every group of order less than 96 the set of commutators form a group! At this point I do not know a way to describe this group.

and then pin it to your fridge so you never make it again! Again, one reason why this is such a common error is that, for one, this is the first time we define an abstract subgroup which is generated by a set, and for two, it is difficult to come up with examples where it's false! However, in any case, it should be clear that to actually find the commutator subgroup by definition, even using the "wrong" definition, would be quite a task!

EXERCISE 11.19. Prove that if $H_1 \preccurlyeq H$ and $K_1 \preccurlyeq K$, then $[H_1, K_1] \preccurlyeq [H, K]$.

SOLUTION

EXERCISE 11.20.

(1) Prove that $[G, H] \preccurlyeq H$ iff H is normal in G.

(2) Use the first part to give a different proof that $[G, G] \trianglelefteq G$ then the one given below.

Hint! It should be obvious that $[G, H] \preccurlyeq [G, G]$ whenever $H \preccurlyeq G$.

SOLUTION

EXERCISE 11.21.

(1) If $\varphi\colon G \to H$ is a homomorphism of groups, then $\varphi([a, b]) = [\varphi(a), \varphi(b)]$.

(2) Use the first part of this exercise to prove that $[G, G]$ **char** G (and hence $[G, G] \trianglelefteq G$).
Hint! Automorphisms are special types of homomorphisms. Make sure you use Corollary 9.28.

(3) Conclude that if $N \trianglelefteq G$, then $[N, N] \trianglelefteq G$.

SOLUTION

EXERCISE 11.22. Prove that both G_{ab} and $[G, G]$ are isomorphism invariants. (Meaning

$$G \cong \widehat{G} \Rightarrow G_{ab} \cong \widehat{G}_{ab} \text{ and } [G, G] \cong \left[\widehat{G}, \widehat{G}\right].)$$

Recall that every subgroup of the center of a group G is normal in G. The following exercise says that the commutator works in the opposite direction in the sense that every supergroup (i.e. superset and group) of the commutator of G is normal in G.

SOLUTION

EXERCISE 11.23. If $[G, G] \preccurlyeq N \preccurlyeq G$, then $N \trianglelefteq G$.
Hint! Use Exercise 11.20**(1)**.

SOLUTION

THEOREM 11.48. *Suppose $N \trianglelefteq G$. Then G/N is abelian iff $[G,G] \leqslant N$.*

PROOF. In a sense, we've proven this already. I give a direct proof of this to make things seem a bit more organized.

$\Rightarrow$ Suppose that G/N is abelian. By definition of the group generated by a set, in order to show that $[G,G] \leqslant N$ it suffices to prove that the commutator of any two elements of G is in N. Indeed, let $a, b \in G$. Then, since G/N is assumed abelian, we have

$$\begin{aligned} aNbN = bNaN &\Rightarrow abN = baN \\ &\Rightarrow (ab)(ba)^{-1} \in N \\ &\Rightarrow aba^{-1}b^{-1} \in N \\ &\Rightarrow [a,b] \in N. \end{aligned}$$

$\Leftarrow$ All implications can be turned around starting from $[a,b] \in N$ and ending with $aNbN = bNaN$. ∎

DEFINITION 11.49. For any group G, the abelianization of G, denoted G_{ab}, is defined as

$$G_{ab} \stackrel{\text{def}}{=} G/[G,G].$$

EXAMPLE /THEOREM 11.50.

$$[\mathcal{S}_n, \mathcal{S}_n] = \mathcal{A}_n.$$

PROOF.

$\subseteq$ Since $\mathcal{S}_n/\mathcal{A}_n \cong \mathbb{Z}_2$ the previous theorem, Theorem 11.48, with $G = \mathcal{S}_n$ and $H = \mathcal{A}_n$, says that

$$[\mathcal{S}_n, \mathcal{S}_n] \leqslant \mathcal{A}_n.$$

$\supseteq$ To prove the other inclusion, note that $\mathcal{S}_{nab} = \mathcal{S}_n/[\mathcal{S}_n, \mathcal{S}_n]$ is abelian (and since $\mathcal{S}_n$ is generated by elements of order 2 ⟨namely the 2- cycles ⟩), and is generated by elements of order 2. As such by the Corollary 9.38 back on page 457, every non-identity element of this quotient has order 2 and so every element whose order is odd in $\mathcal{S}_n$ must be in $[\mathcal{S}_n, \mathcal{S}_n]$.[II] Since 3–cycles have odd order (their order is 3), every 3–cycle must be in $[\mathcal{S}_n, \mathcal{S}_n]$. But the set of m3–cycles generate $\mathcal{A}_n$ and so

$$\text{3–cycles in } \mathcal{S}_n \subseteq [\mathcal{S}_n, \mathcal{S}_n] \Rightarrow \mathcal{A}_n = \left\langle \text{3–cycles in } \mathcal{S}_n \right\rangle \leqslant \langle\, [\mathcal{S}_n, \mathcal{S}_n] \,\rangle = [\mathcal{S}_n, \mathcal{S}_n]$$

as desired. ∎

If $\varphi\colon G \to A$ is a homomorphism from G to an abelian group A, then

$$G/\text{Ker}\,(\varphi) \cong \widetilde{A} \leqslant A$$

[II] This is easy, but may take a short proof.

by the first isomorphism theorem. But since $\widetilde{A}$ is a subgroup of an abelian group, $\widetilde{A}$ is itself abelian. Therefore by the previous theorem this implies that $[G,G] \leqslant \mathrm{Ker}(\varphi)$. This means that we have a well-defined homomorphism

$$G_{ab} \to A$$
$$g[G,G] \mapsto \varphi(g).$$

This is due to Lemma 8.66, but it might be easier for you to prove in this current situation.

Therefore, we have proved that the set of homomorphism from any group G to any abelian group A is the same as the set of homomorphism from the abelianization of G, G_{ab} to A.

THEOREM 11.51. *For any group G and any abelian group A, we have*

$$\hom(G,A) = \hom(G_{ab},A).$$

PROOF. This just summarizes the argument above. ∎

DEFINITION 11.52. A group G is said to be *perfect* if $[G,G] = G$.

THEOREM 11.53. *The following are equivalent:*

(1) *G is perfect;*

(2) *G_{ab} is trivial;*

(3) *Every non-trivial quotient of G is non–abelian;*

(4) *The only homomorphism from G to an abelian group is the constant homomorphism, $g \mapsto e_G$.* ∎

Since for an abelian group A, A' is trivial, $\{e\}$, the only perfect abelian group is the trivial group. Now for a simple group G, by definition, there is no proper, non-trivial normal subgroups. Since $G' \trianglelefteq G$ we must either have $G' = \{e_G\}$ or $G' = G$. The first case means that G is abelian, then second case means that G is perfect. Collecting these facts with some others gives us the following theorem.

THEOREM 11.54.

(1) *A non-trivial, perfect group is non-abelian.*

(2) *A simple, non-abelian group is perfect.*

There are examples of perfect groups which are not simple. The smallest such group is $\mathbf{SL}_2(\mathbb{Z}_5)$ (which is isomorphic to $B\left(\text{★}^3\right)$). (We are about to, in Proposition 11.58, to give a proof of the fact that for all primes $p > 3$, $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ is a perfect group.) Moreover, $\mathbf{SL}_2(\mathbb{Z}_5)$ is not simple since it contains the element $-\mathbf{I}_2$, which, as a scalar matrix, is central and hence generates a non-trivial normal subgroup.

Here is a quick and easy way to see that a group is perfect. (The notation is from Notation 9.37 back on page 456.)

THEOREM 11.55. *Suppose that a group G can be generated by generating sets $X_1 \ldots X_r$ with*

$$\gcd(\mathrm{lcm}(|X_1|),\ldots,\mathrm{lcm}(|X_r|)) = 1.$$

Then G is perfect.

Proof. Let $n_i := \text{lcm}(|X_i|)$. Exercise Exercise 9.19 back on page 457 clearly implies that for all $i \in \overline{r}$

$$G_{ab} = G_{ab}[n_i].$$

But then by Exercise 5.27 back on page 265 we know that

$$G_{ab}[1] = G_{ab}[\gcd(n_1, \ldots, n_r)] = G_{ab}[n_1] \cap \cdots \cap G_{ab}[n_r] = G_{ab} \cap \cdots \cap G_{ab} = G_{ab}.$$

So $G_{ab}[1] = G_{ab}$, i.e. every element of G_{ab} has order 1. But this means that G_{ab} contains only the identity. Therefore G_{ab} is trivial and G is perfect. ∎

Example 11.56. The group $\mathcal{A}_5$ can be generated by 5–cycles (see Exercise 10.75 back on page 590**(1)**) or by 3–cycles (see Exercise 9.40 back on page 477). Since 5 and 3 are relatively prime, Theorem 11.55 implies that $\mathcal{A}_4$ is perfect.

Example 11.57. For all prime numbers, the groups $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ can be generated by elements of order p, $\mathfrak{L}$ and $\mathfrak{R}$ (note $\text{lcm}(|\mathfrak{L}|, |\mathfrak{R}|) = p$), or by an element of order 3, namely Ω and an element of order 4, namely $\mathbf{J}$ from page 445 (note $\text{lcm}(|\Omega|, |\mathbf{J}|) = 12$). For all primes greater than 3, the numbers p and 12 are relatively prime. Therefore, in these cases, Theorem 11.55 applies to show that $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ is perfect. We record this as a proposition:

Proposition 11.58. *For all prime numbers $p > 3$, $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ is perfect.* ∎

Example 11.59. I claim that the group

$$G := \langle (1\,2\,3\,4\,5\,6\,7), (1\,2)(3\,6) \rangle$$

is perfect.

Indeed, we have that

$$(1\,2\,3\,4\,5\,6\,7) \cdot (1\,2)(3\,6) = (1\,3\,7)(4\,5\,6).$$

Therefore each of the following sets generate G:

$$\{(1\,2\,3\,4\,5\,6\,7), (1\,2)(3\,6)\} \quad \{(1\,2\,3\,4\,5\,6\,7), (1\,3\,7)(4\,5\,6)\} \quad \{(1\,2)(3\,6), (1\,3\,7)(4\,5\,6)\}.$$

The least common multiple of the orders are 14, 21, and 6 respectively. But

$$\gcd(14, 21, 6) = \gcd(\gcd(14, 21), 6) = \gcd(7, 6) = 1.$$

Therefore Theorem 11.55 applies to prove that G is perfect.

The generalization of this previous example is clearly:

Theorem 11.60. *Suppose that $G = \langle x, y \rangle$ with $|x|$ and $|y|$ relatively prime. Moreover assume that $|xy|$ is also relatively prime to both $|x|$ and $|y|$. Then G is perfect.*

Exercise 11.24. Prove that $B\left(\text{✶}^3\right)$ is perfect without making use to the non-trivial isomorphism $B\left(\text{✶}^3\right) \cong \mathbf{SL}_2(\mathbb{Z}_5)$.

SOLUTION

Exercise 11.25. Suppose G is perfect and $N \trianglelefteq G$.

(1) Prove that G/N is also perfect.

(2) Is N necessarily perfect?

Solution

11.5.1 Practicing Calculating G_{ab} and $[G, G]$.

Example 11.61. Let G be the group

$$G := \left\langle (1\,2\,3\,\overset{x}{\overset{\shortparallel}{4}}\,5\,6\,7\,8), (1\,5)\overset{y}{\overset{\shortparallel}{(}}3\,7) \right\rangle.$$

Then we have that $|x| = 8$, $|y| = 2$, and $yxy^{-1} = x^5$. This implies that $\langle x \rangle$ is a normal subgroup of order 8 and index 2. In particular, G has order 16. Since x and y do not commute, this is a non–abelian group.

Now since $\langle x \rangle$ is normal in G, we also have that $N := \left\langle x^4 \right\rangle$ is normal in G (it's also easy to check that x^4 is central in G). Then

$$G/N = \left\langle \overset{\bar{x}}{\overset{\shortparallel}{xN}}, \overset{\bar{y}}{\overset{\shortparallel}{yN}} \right\rangle.$$

Moreover, since

$$\overline{yxy}^{-1} = \bar{x}^5 = \cancel{\bar{x}^4} \cdot \bar{x} = \bar{x}$$

we have that $\bar{x}$ and $\bar{y}$ commute, and since they generate the quotient, G/N is abelian. Since the subgroup N contains two elements and since G itself is not abelian, $[G, G] = N$, and $G/N = G_{ab}$.

While we're at it, we may as well point out that $G_{ab} \cong \mathbb{Z}_4 \times \mathbb{Z}_2$ with $\bar{x}$ generating the $\mathbb{Z}_4$ factor and $\bar{y}$ generating the $\mathbb{Z}_2$ factor.

The following exercise would have made the previous example even easier, note that in that example

$$[x, y] = xyx^{-1}y^{-1} = xx^3 = x^4.$$

Exercise 11.26. Suppose $G = \langle S \rangle$. Prove that

$$[G, G] = \lfloor \{[x, x'] : x, x' \in S\} \rceil$$

where $\lfloor\ \rceil$ is the smallest normal subgroup of G which contains the commutators of all generators and is defined on **??**.

Solution

For an application of the previous exercise note that for a dihedral group of degree n, $\mathcal{D}_n$, we have a generating set ρ and $\mathbf{r}$. Moreover $[\rho, \mathbf{r}] = \rho^2$. Since ρ generates a normal subgroup of $\mathcal{D}_n$ so does ρ^2 and so the normal closure of ρ^2 is the same as the usual group generated by ρ^2. As such, we therefore have that $[\mathcal{D}_n, \mathcal{D}_n] = \left\langle \rho^2 \right\rangle$. Now the order of this subgroup is the order of the element generating it, ρ^2 which is

$$\frac{n}{\gcd(2, n)}.$$

This number is n if n is odd and $n/2$ if n is even. Therefore if n is odd, $\langle \rho \rangle = \left\langle \rho^2 \right\rangle$ and $[\mathcal{D}_n, \mathcal{D}_n] = \langle \rho \rangle$ has order n and if n is even, $[\mathcal{D}_n, \mathcal{D}_n] = \left\langle \rho^2 \right\rangle$ has order $n/2$. In the first case, the abelianization is generated just by $\mathbf{r}$ and has order 2 and isomorphic to $\mathbb{Z}_2$. In the second case, the abelianization is generated by ρ and $\mathbf{r}$, both of which have order 2 in the quotient and so in this case the abelianization is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. In summary:

$$(\mathcal{D}_{2k+1})_{ab} \cong \mathbb{Z}_2 \quad \text{and} \quad (\mathcal{D}_{2k})_{ab} \cong \mathbb{Z}_2 \times \mathbb{Z}_2.$$

EXERCISE 11.27. Suppose that $\lfloor S \rceil = G$ for some set $S \subseteq G$. Prove that G_{ab} is generated by S's image.

SOLUTION

EXAMPLE 11.62. As an application of this last exercise we can prove that for $n > 4$, $\mathcal{A}_n$ is perfect. This is implied by the fact that these groups are simple, but this proof is much easier and illustrates a use of the previous exercise, so we go through the motions to prove. Indeed, we always have that

$$\mathcal{A}_n = \lfloor (1\ 2\ 3) \rceil .$$

Therefore, by the previous exercise, $(\mathcal{A}_n)_{ab} = \langle (1\ 2\ 3)_{ab} \rangle$. But for $n > 4$, $(1\ 2\ 3)$ and its inverse $(1\ 3\ 2)$ are conjugate. Actually, the permutation $(1\ 2)(4\ 5)$ does the job. Thus in $(\mathcal{A}_n)_{ab}$,

$$(1\ 2\ 3)_{ab} = (1\ 2\ 3)_{ab}^{-1} \Rightarrow (1\ 2\ 3)_{ab} = (1\ 2\ 3)_{ab}^{-2} = (1)_{ab} .$$

Therefore

$$(\mathcal{A}_n)_{ab} = \langle (1\ 2\ 3)_{ab} \rangle$$

is trivial, and so $\mathcal{A}_n$ is perfect.

11.5.2 More applications

Now since the set of index 2 subgroups of a group G are in bijection with the set of homomorphism $G \to \mathbb{Z}_2$, and the set of such homomorphisms are the same thing as the set of homomorphisms from $G_{ab} \to \mathbb{Z}_2$ we have proven that

THEOREM 11.63.

Index 2–subgroups of $G \overset{1\text{-}1}{\leftrightsquigarrow}$ *non-trivial homomorphisms* $G_{ab} \to \mathbb{Z}_2$ ■

EXAMPLE 11.64. Consider the group

$$G := \left\langle (1\ 2\ 3\ 4\overset{x}{\overset{\shortparallel}{5}}\ 6\ 7\ 8), (1\ 5)\overset{y}{\overset{\shortparallel}{(}}3\ 7) \right\rangle .$$

first define in Example 11.61. In that example, we proved that $G_{ab} \cong \mathbb{Z}_4 \times \mathbb{Z}_2$. Moreover, by Exercise 8.58 back on page 427,

$$\hom(\mathbb{Z}_4 \times \mathbb{Z}_2, \mathbb{Z}_2) \cong \hom(\mathbb{Z}_4, \mathbb{Z}_2) \times \hom(\mathbb{Z}_2, \mathbb{Z}_2) \cong \mathbb{Z}_2 \times \mathbb{Z}_2.$$

Therefore there are 4 homomorphism from G to $\mathbb{Z}_2$ and so there are 3 **non**-trivial ones, and so there are a total of 3 index 2 subgroups. They are the kernels of the maps

$$\begin{array}{lll} \varphi_1(x) = [\,1\,]_2 & \varphi_2(x) = [\,1\,]_2 & \varphi_3(x) = [\,0\,]_2 \\ \varphi_1(y) = [\,0\,]_2 & \varphi_2(y) = [\,1\,]_2 & \varphi_3(y) = [\,1\,]_2 \end{array}$$

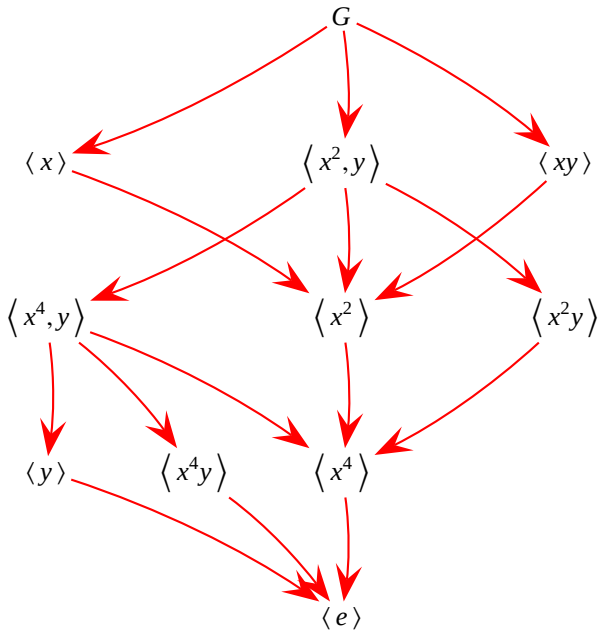
which kernels:

$$K_1 := \left\langle x^2, y \right\rangle = \mathrm{Ker}\,(\varphi_1) \qquad K_2 := \langle\, xy \,\rangle = \mathrm{Ker}\,(\varphi_2) \qquad K_3 := \langle\, x \,\rangle = \mathrm{Ker}\,(\varphi_3)$$

and are in turn isomorphic to

$$K_1 \cong \mathbb{Z}_4 \times \mathbb{Z}_2 \qquad K_2 \cong \mathbb{Z}_8 \qquad K_3 \cong \mathbb{Z}_8.$$

Since every proper subgroup of a 2-group is contained inside of an index 2 subgroup, {why?} this gives us a big jump into constructing the subgroup lattice of G:



11.6 Solvable Groups

Definition 11.65. A *solvability chain* of length m of a group G is a collection of subgroups

$$\{e_G\} = N_0 \trianglelefteq N_1 \cdots \trianglelefteq N_m = G,$$

with

(1) Each group listed is normal in the next;

(2) Each successive quotient, N_{i+1}/N_i, is abelian.

G is *solvable* if it has a solvability chain of finite length. The minimal length of a solvability chain of a solvable group is called the *derived length* of G.

Example 11.66. A solvability chain for an abelian group is

$$\{e_A\} \trianglelefteq A.$$

Hence an abelian group is solvable of derived length 1. Conversely, if G is solvable of derived length 1, then we would have that

$$\{e_G\} \trianglelefteq G$$

is solvability chain for G which means that $G = G/e_G$ itself is abelian. Therefore we have

Groups of derived length 1 $\overset{1\text{-}1}{\leftrightsquigarrow}$ Abelian Groups

Example 11.67. The group $\mathcal{S}_4$ is solvable with solvability chain

$$\{(1)\} \trianglelefteq \langle (1\,2)(3\,4), (1\,3)(2\,4) \rangle \trianglelefteq \mathcal{A}_4 \trianglelefteq \mathcal{S}_4.$$

Example 11.68. The group $B\left(\blacklozenge^3\right)$ is solvable with solvability chain

$$\{1\} \trianglelefteq \langle i \rangle \trianglelefteq Q_8 \trianglelefteq B\left(\blacktriangleright^3\right) \trianglelefteq B\left(\blacklozenge^3\right).$$

Proposition 11.69. *Only the trivial group is solvable and perfect.*

Proof. The assumption G is solvable guarantees that there exists a normal subgroup of G, N, so that G/N is non-trivial and abelian. Indeed, we can take N to be the second to last term in the solvability chain of G. ∎

Example 11.70. None of the groups: $\mathcal{A}_5$, $B\left(\bigstar^3\right)$, or $\mathbf{SL}_2(\mathbb{Z}_7)$ are solvable since we've shown them each to be perfect.

Theorem 11.71.

(1) *If G is solvable, then*

- **(a)** *Every subgroup of G is solvable;*
- **(b)** *Every quotient of G is solvable.*

(2) *If there exists $N \trianglelefteq G$ with N and G/N solvable, then G is solvable.*

PROOF.

(1) Suppose that G is solvable with a solvability chain

$$\{e_G\} = G_0 \leqslant G_1 \leqslant \cdots \leqslant G_m = G.$$

- **(a)** Suppose that $H \leqslant G$.
  Claim The chain

$$\{e_H\} = \overset{H_0}{\overset{\shortparallel}{H \cap G_0}} \leqslant \cdots \leqslant \overset{H_i}{\overset{\shortparallel}{H \cap G_i}} \leqslant \cdots \leqslant \overset{H_m}{\overset{\shortparallel}{H \cap G_m}} = H \tag{4}$$

  is a solvability chain for H.

  By definition, all we need to do is show that each successive quotient, H_{i+1}/H_i is abelian. We note that since $H_i \leqslant H_{i+1}$ we have $H_i = H_{i+1} \cap G_i = H \cap G_i$. Indeed, by second isomorphism theorem, each quotient

$$\begin{aligned} H_{i+1}/H_i &= H_{i+1}/H \cap G_i && \text{(Def'n of } H_i\text{)} \\ &= H_{i+1}/H_{i+1} \cap G_i && \text{(Using } H_{i+1} \cap G_i = H \cap G_i\text{)} \\ &\cong H_{i+1} \cdot G_i/G_i && \text{(Using } 2^{\text{nd}} \text{ iso theorem)} \\ &\leqslant G_{i+1}/G_i. && (H_{i+1}, G_i \leqslant G_{i+1} \Rightarrow H_{i+1} \cdot G_i \leqslant G_{i+1}) \end{aligned}$$

  But this last quotient, G_{i+1}/G_i is abelian. Therefore, the quotient that we started with, H_{i+1}/H_i, as a subgroup of an abelian group, is also an abelian group. Putting everything together guarantees us that (4) is a solvability chain for H.

- **(b)** Suppose $N \trianglelefteq G$.
  Claim The chain

$$\{e_{G/N}\} = \underbrace{(N \cdot G_0)/N}_{\shortparallel\, N_0} \leqslant \cdots \leqslant \underbrace{(N \cdot G_i)/N}_{\shortparallel\, N_i} \leqslant \cdots \leqslant \underbrace{(N \cdot G_m)/N}_{\shortparallel\, N_m} = G/N \tag{5}$$

  is a solvability chain for the quotient, G/N.

  By definition, all we need to do is show that each sucessive quotient, N_{i+1}/N_i is abelian. But these quotients are

$$\begin{aligned} N_{i+1}/N_i &= N \cdot G_{i+1}/N / N \cdot G_i/N \\ &= N \cdot G_{i+1}/N \cdot G_i && \text{(Using } 3^{\text{rd}} \text{ Isomorphism Theorem)} \end{aligned}$$

(2) Suppose that

$$\{e_{G/N}\} = N_0/N \trianglelefteq N_1/N \cdots \trianglelefteq N_m/N = G/N.$$

is a solvability chain for G/N and that

$$\{e_H\} = N_0' \trianglelefteq N_1' \trianglelefteq \cdots \trianglelefteq N_k' = N$$

is a solvability chain for N. Then I claim that

$$\{e_H\} = N_0' \trianglelefteq N_1' \trianglelefteq \cdots \trianglelefteq N_{k-1}' \trianglelefteq \overset{N_0=N_k'}{\overset{\shortparallel}{N}} \trianglelefteq N_1 \trianglelefteq \cdots \trianglelefteq N_m = G$$

is a solvability chain for G. Note that the first k quotients are of the form N_i'/N_{i-1}', which are abelian by assumption, (they are the i^{th} quotient of the solvability chain for N). Meanwhile the last m quotients are of the form

$$N_j/N_{j-1} \cong (N_j/N)/(N_{j-1}/N)$$

by the 3rd Isomorphism Theorem. Since this is the j^{th} quotient in the solvability chain for G/N this quotient is also abelian. ∎

EXERCISE 11.28. Suppose that H and K are perfect groups which happen to be subgroups of some group. Prove that

$$\langle H \cup K \rangle$$

is a perfect group.
Hint! Consider a homomorphism

$$\varphi\colon \langle H \cup K \rangle \to A$$

where A is some abelian group. Prove that $\varphi|_H$ and $\varphi|_K$ are trivial, then conclude that φ is also trivial.

SOLUTION

EXERCISE 11.29. Let G be a finite group. Prove that

(1) G contains a unique maximal, perfect subgroup, G^{perf};

(2) G^{perf} is characteristic in G.

(3) G is solvable iff $G^{perf} = \{e_G\}$;

(4) G/G^{perf} is solvable.

SOLUTION

EXAMPLE 11.72. For $n \geq 5$, $\mathcal{S}_n$ is not solvable. This is because all such groups contain the non-solvable group $\mathcal{A}_5$ as a subgroup. (See Example 11.56 for one proof of the fact that $\mathcal{A}_5$ is not solvable.)

Here is a more computable way of telling whether or not a group is solvable.

DEFINITION 11.73. The *derived series* of a group is defined inductively as follows:

(1) $\langle [\,\rangle 0] \stackrel{\text{def}}{=} G$.

(2) Assuming that $G^{(i)}$ has been defined

$$G^{(i+1)} \stackrel{\text{def}}{=} \left[G^{(i)}, G^{(i)}\right].$$

Therefore $G^{(1)} = [G,G]$, $G^{(2)} = [[G,G],[G,G]]$ etc. Since the commutator of a group K, $[K,K]$, is characteristic in K we know that each term of the derived series is characteristic in the previous. But this implies that each term in the derived series in normal in G. Moreover, since each term of the derived series in the commutator of the previous, we know that each quotient

$$G^{(i)}/G^{(i+1)}$$

is abelian. Therefore, if there is some integer n with $G^{(n)} = e_G$ the derived series forms a solvability chain for G and as such G is solvable.

On the other hand, suppose G has a solvability chain

$$e = N_0 \trianglelefteq N_1 \trianglelefteq \cdots \trianglelefteq N_m = G.$$

Then since G/N_{m-1} is abelian, we have that $[G,G] \trianglelefteq N_{m-1}$. Since N_{m-1}/N_{m-2} is also abelian, we know that $[N_{m-1}, N_{m-1}] \trianglelefteq N_{m-2}$. Coupling this with the fact that $[G,G] \trianglelefteq N_{m-1}$ we see that

$$G^{(2)} = [[G,G],[G,G]] \trianglelefteq [N_{m-1}, N_{m-1}] \trianglelefteq N_{m-2}.$$

Continuing like this, we can see that $G^{(i)} \trianglelefteq N_{m-i}$. As such, since $N_0 = e$, we know that $G^{(m)}$ is contained inside of the trivial group, and so it must also be trivial. Moreover, in this case, this also implies that the derived series is a solvability chain of shortest length. Collecting everything:

THEOREM 11.74. *A group G is solvable iff there exists an integer m so that $G^{(m)} = e_G$. Moreover, G has derived length d iff the least integer where this holds if d.*

Note also that if G happens to be finite, the chain

$$G^{(m+1)} \trianglelefteq G^{(m)} \trianglelefteq G^{(m-1)}$$

can't keep getting smaller forever. Therefore there exists some integer M where the chain stops, i.e.

$$\cdots G^{(M+2)} \trianglelefteq G^{(M+1)} \trianglelefteq G^{(M)}.$$

In this case, G is solvable iff this stopping point is the trivial subgroup, $\{e_G\}$.

EXAMPLE 11.75. (This is a more straight forward way of reproducing a result that's a direct consequence of Example 11.23 back on page 614)

The two groups $Q_8 \times \mathbb{Z}_2$ and ${}_4G := \left\langle (1\ \overset{x}{2}\ 3\ 4), (1\ 3)(\overset{y}{5}\ 6\ 7\ 8) \right\rangle$ from Example 11.11 are two non-abelian groups of order 16. Moreover a calculation shows that they both have order tables

n	1	2	4	8	16
$\mathbf{o}_n$	1	3	12	0	0

To make things seemingly worse, both have a center which is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

This seems to exhaust all of our usual tricks of proving that these two groups are not isomorphic! However, armed with abelianizations, we can see that the abelianization of $Q_8 \times \mathbb{Z}_2$ is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$ (with derived subgroup $\langle (-1,0),(0,1) \rangle$). On the other hand, we showed in the example just mentioned that $G/\left\langle x^2 \right\rangle$ is abelian (and

isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_4$.) Since ${}_4G/\left\langle x^2 \right\rangle$ is abelian, we know that $[G,G] \preccurlyeq \left\langle x^2 \right\rangle$. But since $\left\langle x^2 \right\rangle$ only contains the trivial subgroup as a proper subgroup, and since $G = G/e$ is non-abelian, we know that $[G,G] = \left\langle x^2 \right\rangle$ and

$$G_{ab} \cong \mathbb{Z}_2 \times \mathbb{Z}_4.$$

Therefore ${}_4G$ and $Q_8 \times \mathbb{Z}_2$ have non-isomorphic abelianizations and are therefore not isomorphic to each other.

EXERCISE 11.30. Prove that a finite group is not solvable iff it contains a non-trivial perfect subgroup.
Hint! Consider the term in the derived series where it "stops getting smaller."

SOLUTION

EXERCISE 11.31. **T** or **F**:

SOLUTION

(1) The intersection of two perfect groups is perfect.

(2) The intersection of two solvable groups is solvable.

(3) Every solvable group is abelian.

(4) Every cyclic group is solvable.

(5) Every simple group is perfect.

(6) Every perfect group is solvable.

(7) The group generated by two perfect subgroups of a common group is a perfect subgroup.

(8) The group generated by two solvable subgroups of a common group is a perfect subgroup.

11.7 APPLICATIONS TO P–GROUPS.

In this section we use quotient groups to classify groups of order p^2 as being either isomorphic to $\mathbb{Z}_{p^2}$ or $\mathbb{Z}_p \times \mathbb{Z}_p$ (as usual, you can verify that all groups of orders

$$2^2 = 4, 3^2 = 9, 5^2 = 25, 7^2 = 49, 11^2 = 121$$

are abelian by checking out Appendix I which starts on page 865). First a lemma which is of independent interest.

LEMMA 11.76. *Suppose $N \preccurlyeq \mathcal{C}(G)$. If G/N is cyclic, then G is abelian.*

PROOF. Suppose that G/N is cyclic with generator tN. This means for all $g \in G$ there exists an integer i so that $gN = t^i N$. As such, for arbitrary $g, h \in G$ there exists i and j integers and z_1 and z_2 central with

$$g = t^i z_1 \quad \text{and} \quad h = t^j z_2. \tag{6}$$

Therefore we have

$$\begin{aligned} gh &= t^i z_1 t^j z_2 && \text{(Plugging in via Equation (6))} \\ &= t^i t^j z_1 z_2 && \text{(Since } z_1 \in \mathcal{C}(G)\text{)} \\ &= t^j t^i z_1 z_2 && \text{(Common powers of } t \text{ com)} \\ &= t^j t^i z_2 z_1 && \text{(Since } z_1 \in \mathcal{C}(G)\text{)} \\ &= t^j z_2 t^i z_1 && \text{(Since } z_2 \in \mathcal{C}(G)\text{)} \\ &= hg. \end{aligned}$$

Therefore g and h commute. Since they were arbitrary elements of G, G is abelian. ■

It is worth noticing that while the statement of the above theorem requires quotient groups, its proof does not (it only uses the properties of cosets).

EXERCISE 11.32. Show that for any group G that TFAE

- $G/\mathcal{C}(G)$ is cyclic.
- $G/\mathcal{C}(G)$ is trivial.
- G is abelian.

SOLUTION

This means if you do some calculation and prove that $G/\mathcal{C}(G)$ is (say) isomorphic to $\mathbb{Z}_4$, you did something wrong!

EXERCISE 11.33. Prove the for any group G, $\left|G/\mathcal{C}(G)\right|$ is not prime.

SOLUTION

EXERCISE 11.34. Find a non-abelian group G with $G/\mathcal{C}(G)$ is abelian. (The theorem only rules out the case that $G/\mathcal{C}(G)$ is cyclic. Of course there are a ton of groups which are abelian but which are not cyclic. For instance, $\mathbb{Z}_2^2$.)

SOLUTION

EXERCISE 11.35. Show that a non-abelian group of order 16 has center either of order 2 or order 4. (Just eliminate 1, 8 and 16!)

SOLUTION

EXERCISE 11.36. If a group G of order 16 has center of order 4, what must the quotient $G/\mathcal{C}(G)$ be isomorphic to?

SOLUTION

EXERCISE 11.37. Suppose G is a group of order p^4 (p prime) and that $Z \preccurlyeq \mathcal{C}(G)$ has order p^2. Prove that $Z = \mathcal{C}(G)$.

SOLUTION

EXERCISE 11.38. Suppose G is a non–abelian group of order p^3. Prove that $[G,G] = \mathcal{C}(G)$ and both have order p.

SOLUTION

EXERCISE 11.39. Suppose G is a group of order p^4. Prove that at least one of the following are true: $[G,G] \subseteq \mathcal{C}(G)$ or $\mathcal{C}(G) \subseteq [G,G]$.
Hint! First deal with trivial (or impossible cases) that $|\mathcal{C}(G)| = p^4$ or $|\mathcal{C}(G)| = 1$. Next, If $|\mathcal{C}(G)| = p^2$, conclude that $[G,G] \subseteq \mathcal{C}(G)$. Lastly, if $|\mathcal{C}(G)| = p$, then use Exercise 10.68 back on page 583 to conclude that $\mathcal{C}(G) \subseteq [G,G]$.

SOLUTION

EXAMPLE 11.77. Let

$$G = \left\langle (1\,\overset{x}{2}\,3\,4), (1\,3)(\overset{y}{5}\,6\,7\,8) \right\rangle.$$

be the group from Example 11.11. We have seen that this group has order 16. A direct calculation shows that $yxy^{-1} = x^{-1}$. Therefore, as x has order 4, $yx^2y^{-1} = x^{-2} = x^2$ and

$$y^2xy^{-2} = yx^{-1}y^{-1} = \left(yxy^{-1}\right)^{-1} = (x^{-1})^{-1} = x.$$

As such both x^2 and y^2 are central (since they commute with both generators). Moreover, they are distinct and each have order 2. Therefore $\left\langle x^2, y^2 \right\rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$, in particular this subgroup has order 4. Since the center of a non-abelian group of order 16 can not have more than 4 elements, we see that

$$\mathcal{C}(G) = \left\langle x^2, y^2 \right\rangle.$$

EXERCISE 11.40. Verify this last example by drawing the conjugation diagram for this group.

SOLUTION

EXERCISE 11.41. Let p and q be distinct prime numbers. Show that a non-abelian group of order pq is centerless.
Hint! Check out the proof of Theorem 11.78 just below! (Of course examples of such groups are $\mathcal{D}_5$, $\mathcal{D}_7$, or more generally $\mathcal{D}_p$ where p is an odd prime integer.)

SOLUTION

THEOREM 11.78. *Any group of order p^2 is abelian (p prime).*

Proof. Suppose that G contradicts the statement of the theorem. That is assume that G is a non–abelian group of order p^2. Then, we know that $\mathcal{C}(G)$ is a proper subgroup of G (since G is assumed non-abelian, $G \neq \mathcal{C}(G)$), which in addition is also non-trivial (its non-trivial since p–groups are never centerless by Theorem 10.123). Hence, since the only divisors of p^2 are 1, p, and p^2, $\mathcal{C}(G)$ must have order p. Therefore $G/\mathcal{C}(G)$ must have order $p = p^2/p$. But any group of order p is cyclic, therefore Lemma 11.76 applies to prove that G is abelian. ■

The following illustrates another way to use quotient groups and the fact that $\mathfrak{p}$–groups have a non-trivial center.

Example /Theorem 11.79. Let G be a p–group. (So the order G is p^n for some positive integer n.) Then for every divisor p^a of $|G|$ there exists a normal subgroup of G of order p^a.

Proof. Instead of proving the theorem as stated, we prove:

> For every $0 \leq a \leq n$, G contains a normal subgroup of index p^a.

This is because the set of divisors of a number are the same thing as the set of quotients of that number by its divisors.

We prove this by induction on n.

- If $n = 1$, then $G \cong \mathbb{Z}_p$. The result clearly holds in this special case.
- Assume the result holds for the case $n - 1$. That is assume that every p–group of order p^{n-1} contains a normal subgroup of order any given divisor of p^{n-1}.

  Suppose G is a group of order p^n. Since G is a p–group, we have that $\mathcal{C}(G)$ is a non-trivial (Theorem 10.123) subgroup of G. As such, $p \Big| |\mathcal{C}(G)|$, and so there exists an element in $\mathcal{C}(G)$, call it g, of order p. This is Cauchy's lemma applied to the group $\mathcal{C}(G)$.

  Since $g \in \mathcal{C}(G)$, $\langle g \rangle \trianglelefteq \mathcal{C}(G)$, and like every subgroup of the center, is a normal subgroup of G. Since g has order p, $\langle g \rangle$ also has order p. Hence $G/\langle g \rangle$ is a group of order $p^n/p = p^{n-1}$. Therefore, the induction hypothesis applies to this quotient group $G/\langle g \rangle$, implying that it (the quotient) has a normal subgroup of index p^a for all $a \leq n - 1$. But the inverse image of this subgroup under the quotient map

  $$\pi : G \to {}^{G}\!/_{\langle g \rangle}$$

  is also a normal subgroup of index p^a, but this time it lives in G. Thus we have proven for all $a \leq n - 1$, G has a normal subgroup of index p^a. The proof is complete since G also has a normal subgroup of index p^n, namely the trivial subgroup. Thus we have proven that for $a \leq n$, G contains a normal subgroup of index a, completing the inductive step and hence the proof. ■

Exercise 11.42. . Prove that every element of a non-cyclic p–group is contained inside of a subgroup of index p.
Hint! Induction again. Select a central element of order p and take the quotient to use the inductive step.

Solution

Exercise 11.43. Prove that for any proper subgroup of a p–group, $H \trianglelefteq P$, we have that $N_P(H) \neq H$.
Hint! A fancy proof would use **OPST** applied to the action $H \circlearrowleft (P/H)^{\times}$ and Exercise 10.34 back on page 530. A less fancy proof would make use of a careful induction. ($(P/H)^{\times}$ denotes $P/H \trianglelefteq \{H\}$.)

Solution

Exercise 11.44. Every index p–subgroup of a P group is normal.
Hint! We will prove this differently (and more generally) a bit later. For now, you can do this in one line by making use of the previous exercise, Exercise 11.43.

Solution

The next exercise seems stronger than 11.42.

Exercise 11.45. Prove that every proper subgroup of a p–group is contained inside of an index p-subgroup.
Hint! Induction again, but this time on the index of the subgroup. Use 11.43 and pull back a nontrivial subgroup of the quotient of the subgroup in its normalizer.

Solution

The example/theorem above guarantees us that index p–subgroups of any p–group exists. Therefore, for a p–group P, it makes sense to define the intersection of all index p subgroups of P. We denote this intersection has $\mathfrak{G}(P)$. It is known as the *Frattini subgroup of* P. In the above exercises, you proved that every subgroup of index p in a p–group is normal. Therefore, as the intersection of normal subgroup of P, $\mathfrak{G}(P) \trianglelefteq P$. Actually, looking at this from another point of view, since an automorphism of a group takes index p–subgroups to index p–subgroups, we actually see that $\mathfrak{G}(P)$ **char** P, which as you should recall, is even stronger than simply saying that it's normal in P.

Next, note that for any index p–subgroup of P, the quotient, having order p, is cyclic and hence abelian. This means that the commutator $[P, P] \leqslant \mathfrak{G}(P)$ Furthermore, since each of this quotients are isomorphic to $\mathbb{Z}_p$, every $g \in G$, g^p is in every index p–subgroup. These facts clearly mean that the quotient, $P/\mathfrak{G}(P)$ is an elementary abelian group, i.e. it is isomorphic to $\mathbb{Z}_p^r$ for some integer r.

Notation 11.80. The *rank* of a P group in the integer r.

Example 11.81. Since cyclic groups $\mathbb{Z}_{p^n}$ have a unique subgroup of index p, which is $\left\langle [\, p \,]_{p^n} \right\rangle$, this must be the Frattini subgroup of $\mathbb{Z}_{p^n}$. More generally, the Frattini subgroup of the abelian p–group

$$\mathbb{Z}_{p^{n_1}} \times \cdots \times \mathbb{Z}_{p^{n_m}}$$

can easily be seen to be

$$\underbrace{\mathbb{Z}_p \times \cdots \times \mathbb{Z}_p}_{m-\text{copies}}$$

and generated by

$$(0, \ldots, 0, [\, 1 \,]_{p^{n_i}}, 0 \ldots, 0).$$

Lemma 11.82. *$g \in \mathfrak{G}(P)$ iff for every proper subgroup H, $\langle x, H \rangle$ is also a proper subgroup of P.*

Proof.

$\Rightarrow$ Suppose that $g \in \mathfrak{G}(P)$ and let H be a proper subgroup of P. By Exercise 11.45 H is contained inside of an index p– subgroup P' of P. But by definiton of $\mathfrak{G}(P)$ we know that g is also contained inside of P'. Hence $\langle g, H \rangle \subseteq P' \prec$.

⇐ We prove the converse by contraposition. Suppose $g \notin \mathfrak{G}(P)$. Then there exists an index p–subgroup of P. Call this subgroup P'. Since $g \notin P'$, we know that $P' \prec \langle g, P' \rangle$. But since P' has index p, this only leaves the possibility that $\langle g, P' \rangle = P$. ∎

COROLLARY 11.83. *If X generates a p–group P, then so does $X \setminus \mathfrak{G}(P)$.*

PROOF. Suppose that $g \in \mathfrak{G}(P) \cap X$ and assume that $X' := X \setminus \{x\}$ is not a generating set for P, i.e. $\langle X' \rangle \neq P$. But by the lemma, we also know that $\langle g, X' \rangle \neq P$. But

$$\langle g, X' \rangle = \langle \{x\} \cup X' \rangle = \langle X \rangle = P$$

a contradiction. ∎

THEOREM 11.84. *The smallest possible number of generators of a p–group P is its rank,* $\mathrm{rk}(P)$.

PROOF. Let's introduce the temporary notation $\mathrm{rk}'(P)$ to mean smallest possible number of generators P can have. Indeed, that Frattini subgroup of such a group is trivial.

We wish to prove that $\mathrm{rk}(P) = \mathrm{rk}'(P)$. Let's denote the lhs as r and the rhs as r'. Clearly the rank of an elementary abelian group $\mathbb{Z}_p^s$ is s. Moreover since the group generated by s elements of $\mathbb{Z}_p^s$ can have at most p^s elements, we have that $\mathrm{rk}(P) = \mathrm{rk}'(P)$ in this case. Since P maps onto $\mathbb{Z}_p^r$, and any generating set of P maps onto a generating set of P^r we know that $r \leq r'$.

On the other hand, suppose $X = \{x_1, \ldots, x_s\}$ is a generating set of P. (Note that by definition, $s \leq r'$.) Assume that $r \leq s$. This means there exist positive integers $n_i \in \mathbb{Z}$ each strictly less than p and at least one non-zero, so that

$$g := x_1^{n_1} \cdots x_s^{n_s} \in \mathfrak{G}(P).$$

WLOG we assume that the non–zero coefficient is n_1. Then as an integer strictly less than p and greater than 0, n_1 is relatively prime with p and hence also to x_1 (since we are in a p–group, $|\, x_1 \,|$ is a power of p). Solving for $x_1^{n_1}$ and using the fact that $\gcd(n_1, |\, x_1 \,|) = 1$ implies that

$$x_1 \in \langle g, x_2, \ldots, x_s \rangle.$$

Therefore, $\langle g, x_2, \ldots, x_s \rangle$ contains every element of X and it is also a generating set of P. But since $g \in \mathfrak{G}(P)$ we can apply the corollary just above, Corollary 11.83, to conclude that $\langle x_2, \ldots, x_s \rangle$ is also a generating set of P. Therefore the assumption that $r < s$ implies that $r' < s$ as well. Hence (taking $s = r + 1$) we have proven the other equality, $r' \leq r$ and so $r = r'$ as desired. ∎

11.8 MORE

THEOREM 11.85. *The alternating group of degree n, $\mathcal{A}_n$, is the only index two subgroup of $\mathcal{S}_n$.*

PROOF. Since $\mathcal{S}_2 \cong \mathbb{Z}_2$, the theorem is obvious for $n = 2$. Therefore we assume that $n > 2$. Let N be an index two subgroup of $\mathcal{S}_n$. Since N is index 2, Theorem 10.61 applies to show that $N \trianglelefteq \mathcal{S}_n$. Since there is only one group of order 2, $\mathcal{S}_n / N \cong \mathbb{Z}_2$. Since all three cycles are conjugate, and $\mathcal{A}_n$ is generated by $\underline{3}$–cycles, it suffices to show that a $\underline{3}$–cycle is contained in N. Now since N is a proper normal subgroup of $\mathcal{S}_n$, N cannot contain a $\underline{2}$–cycle. As such N cannot contain either of the $\underline{2}$–cycles $(1\ 2)$ or $(2\ 3)$. However since the quotient is isomorphic to $\mathbb{Z}_2$ we must have that

$$(1\ 2)(2\ 3)N = ((1\ 2)N)((2\ 3)N) = N \text{ or } (1\ 2\ 3) \in N$$

as desired. ∎

PROPOSITION 11.86. *$\mathcal{A}_n$ never contains an index 2 subgroup.*

PROOF. We know that $\mathcal{A}_n$ is generated by $\underline{3}$–cycles. Therefore if N is an index 2 subgroup of $\mathcal{A}_n$ we must have that the quotient $\mathcal{A}_n / N$ is generated by the image of the $\underline{3}$–cycles in the quotient. But the image of an element of order 3 under any homomorphism has order 1 or 3 depending on whether or not the element is in the kernel or not. If any $\underline{3}$–cycles is in the kernel N, then N would have to be equal to all of $\mathcal{A}_n$ (by a previous homework assignment). In particular, the image of a $\underline{3}$–cycles has order 3 in the quotient, but the quotient is isomorphic to $\mathbb{Z}_2$ which does not have an element of order 3, a contradiction. ∎

The following exercise encompasses what we actually used to prove the above theorem and can be taken as a generalization of it.

EXERCISE 11.46. Suppose a group G is generated by elements of odd order. Show that G contains no index 2 subgroup.

SOLUTION

COROLLARY 11.87. *Any 4–cycle and any 3–cycles in $\mathcal{S}_4$ generates $\mathcal{S}_4$.*

PROOF. Let H be the group generated by the first and K be the group generated by the second. Since 4 and 3 are relatively prime, we have that HK has 12 elements in it and is contained inside of the group G generated by the given 4 and 3–cycles. Therefore the group generated by the two cycles has at least 12 elements. Therefore it has either 12 or 24 elements in it by Lagrange's theorem. It cannot have 12 elements since if it did it would have to be $\mathcal{A}_4$ by the previous theorem. But $\mathcal{A}_4$ does not contain any 4–cycles since such cycles are odd. Therefore $G = \mathcal{S}_4$. ■

EXAMPLE 11.88. Each of $(1\,2\,3), (1\,2\,3\,4)$, $(1\,3\,2), (1\,4\,2\,3)$, and $(1\,2\,3), (1\,2\,4\,3)$ therefore generated $\mathcal{S}_4$. In fact, since there are a total of 8 3–cycles and 6 4–cycles in $\mathcal{S}_4$ there are therefore 48 generating sets of this form!

11.8.1 RECALCULATING $\mathbf{O}_a(\mathbb{Z}_n)$.

In this section we give a much faster proof of the fact that

$$\mathbf{O}_a(\mathbb{Z}_n) = \gcd(n, a)$$

then what was given in 5.89 back on page 294 in Chapter 5. For an integer $a \in \mathbb{Z}$, and a group G recall the notations

$$G[\,a\,] \stackrel{\text{def}}{=} \{g \in G : g^a = e_G\} \quad \text{and} \quad \mathbf{O}_a(G) \stackrel{\text{def}}{=} \#\left(G[\,a\,]\right).$$

PROPOSITION 11.89.

(1) $\mathbb{Z}_n[\,a\,] \leqslant \mathbb{Z}_n$.

(2) $|\,\mathbb{Z}_n[\,a\,]\,| = \gcd(n, a)$

PROOF. Define a map

$$\begin{aligned} \mu_a &: \mathbb{Z}_n \to \mathbb{Z}_n \\ &: [\,m\,]_n \mapsto a\,[\,m\,]_n . \end{aligned}$$

The distributive property in $\mathbb{Z}_n$says that this map is a homomorphism. By definition, the kernel of this homomorphism is $\mathbb{Z}_n[\,a\,]$. This proves (1) since the kernel of any homomorphism is a subgroup of the homomorphism's domain. The image of this homomorphism is $a\mathbb{Z}_n$. But we've seen that $a\mathbb{Z}_n = \gcd(n, a)\mathbb{Z}_n$ and so

$$|\,a\mathbb{Z}_n\,| = \frac{n}{\gcd(n, a)}.$$

Therefore by KIT, we know that

$$|\,\mathbb{Z}_n[\,a\,]\,| = \frac{|\,\mathbb{Z}_n\,|}{|\,a\mathbb{Z}_n\,|} = \frac{n}{n/\gcd(n, a)} = \gcd(n, a)$$

which proves (2). as desired. ■

COROLLARY 11.90. *For any integer a, $\mathbf{O}_a(n) = \gcd(a, n)$.*

PROOF. Part (2) of the proposition says that

$$\gcd(a, n) = |\mathbb{Z}_a|.$$

But by its very definition, $\mathbf{O}_a(\mathbb{Z}_n) = |\mathbb{Z}_n[a]|$, and so making this substitution the result is proven. ■

12 Semi-Direct Products

As originally pointed out, direct product are the easiest way to taking two groups, sticking them together, and forming another group. We have already seen several groups turn out to be direct products of smaller groups that we certainly didn't expect. For instance, we've seen that $\mathbb{Z}_6 \cong \mathbb{Z}_2 \times \mathbb{Z}_3$.

Direct products are fairly rare in group theory. On the other hand, semi-direct products are quite common. For instance, for every prime other than 2, every group of order p^3 or p^4 are (non-trivial) semi-direct products. Every group of order pq for p and q primes, with the exception of $p = q$ and group $\mathbb{Z}_{p^2}$ is a (non-trivial) semi–direct product of groups, and $2p^2$, with the exception of Q_8, is a (non-trivial) semi–direct products.

12.1 More on Direct Products

Before moving onto Semi–Direct Products, I spend a few minutes looking more at their simpler baby brothers: direct products.

Exercise 12.1. The obvious generalization of the direct product recognition theorem is false. The obvious generalization would state (N_1, N_2, and N_3 subgroups of G)

- The N_i's have mutually trivial intersections. That is
$$N_1 \cap N_2 = \{e_G\} \qquad N_1 \cap N_3 = \{e_G\} \qquad N_2 \cap N_3 = \{e_G\}.$$
- Each of the N_i's are normal in G. That is
$$N_1 \trianglelefteq G \qquad N_2 \trianglelefteq G \qquad N_3 \trianglelefteq G.$$
- Together the N_i's generate G. That is
$$\langle N_1 \cup N_2 \cup N_3 \rangle = G.$$

We'd hope that the conclusion would be that $G \cong N_1 \times N_2 \times N_3$. Show, however, that this is false.
Hint! Consider $G = \mathbb{Z}_2 \times \mathbb{Z}_2$. Show that
$$N_1 := \langle (1,0) \rangle \qquad N_2 := \langle (0,1) \rangle \qquad N_3 := \langle (1,1) \rangle$$
satisfy all of the bullet points above, but
$$\mathbb{Z}_2 \times \mathbb{Z}_2 \ncong N_1 \times N_2 \times N_3.$$
(How many elements are in the latter set?)

Solution

Exercise 12.2. Use the same notation as the previous exercise. Show that if $N_1 \cap N_2N_3$ is trivial, then we can safely conclude that

$$G \cong N_1 \times N_2 \times N_3.$$

Solution

Example 12.1.

(1) In this example I prove that $\mathcal{D}_6 \cong \mathcal{D}_3 \times \mathbb{Z}_2$. Let ρ and $\mathbf{r}$ have their usual meanings in $\mathcal{D}_6$ (a rotation by $\pi/3$ and $\mathbf{r}$ a reflection). For starters, since

$$\mathbf{r}\rho^3\mathbf{r}^{-1} = \rho^{-3} = \rho^3$$

we have that $\rho^3 \in \mathcal{C}(\mathcal{D}_6)$ and as such

$$H := \{\mathbf{1}, \rho^3\} = \left\langle \rho^3 \right\rangle \trianglelefteq \mathcal{D}_6.$$

Since this group has two elements in it, it is isomorphic to $\mathbb{Z}_2$. Next, if we look at $K := \left\langle \rho^2, \mathbf{r} \right\rangle$, then if we take $A = \rho^2$ and $B = \mathbf{r}$ we see that $A^3 = \mathbf{1}$, $B^2 = \mathbf{1}$ and $BAB^{-1} = A^{-1}$. Hence this group is isomorphic to $\mathcal{D}_3$.[I] This K is normal since we know that it has index 2. Lastly you can check that H and K have trivial intersection in a number of ways. For example, one can easily write down all of the elements of each subgroup. A fancier way is to simply note that if $\rho^3 \in K$ then

$$\rho^3 \in K \Rightarrow \rho^3\rho^{-2} = \rho \in K.$$

Therefore since both ρ and $\mathbf{r}$ are in K and K would have to be all of $\mathcal{D}_6$ (since it would contain both generators of this group), which it isn't since $\mathcal{D}_6$ has order 12 while K only has order 6. Therefore we have that

$$\mathcal{D}_6 = H \times K \cong \mathbb{Z}_2 \times \mathcal{D}_3$$

as desired.

(2) Let $S^*(\blacklozenge^3)$ be the group of all symmetries of the cube (including the reflections). I claim that we have

$$S^*(\blacklozenge^3) \cong S(\blacklozenge^3) \times \mathbb{Z}_2$$

where (as usual) $S(\blacklozenge^3)$ is the group of rotations of the cube.

- The function $-\mathbf{I}_3 \colon (x, y, z) \mapsto (-x, -y, -z)$ is a symmetry of the cube (it preserves all vertices). Moreover, it is clearly central since it commutes with all linear maps $\mathbf{A}$;

  $$\mathbf{A}([-\mathbf{I}_3](\mathbf{v})) = \mathbf{A}(-\mathbf{v}) = -\mathbf{A}(v) = [-\mathbf{I}_3](\mathbf{A}(v))$$

  and the symmetries of the cube are linear. Obviously the element $-\mathbf{I}_3$ has order two, and so $\langle -\mathbf{I}_3 \rangle = \{\mathbf{I}_3, -\mathbf{I}_3\}$. Next, we know that $S(\blacklozenge^3) \trianglelefteq S^*(\blacklozenge^3)$ since the former has index 2 inside of $S^*(\blacklozenge^3)$.[II]

[I] You can actually see this isomorphism by looking at an inscribed triangle inside of a hexagon and noticing that this subgroup is the set of symmetries of this triangle!

[II] (it's actually the kernel of the homomorphism

$$\det \colon S^*(\blacklozenge^3) \to \{-1, +1\}.$$

- Next
$$\left\langle S(\blacklozenge^3), -\mathbf{1}_{\mathbb{R}^3} \right\rangle = S^*(\blacklozenge^3)$$
by a Lagrange's theorem argument.
- Lastly, the two groups intersect trivially since the (lone) non-trivial element of the first reverses orientation while every rotation preserves it.

The proof it complete since the first group is isomorphic to $\mathbb{Z}_2$.

EXAMPLE 12.2. If $\langle X, Y \rangle \subseteq \mathcal{S}_n$ has that the support of X is disjoint from the support of Y then it is easy to check that
$$\langle X, Y \rangle \cong \langle X \rangle \times \langle Y \rangle .$$
by using the recognition theorem (check each of the three requirements). For instance
$$\langle (1\,2\,3\,4), (5\,7), (5\,6\,7), (1\,3) \rangle$$
satisfies this hypothesis with
$$X = \{(1\,2\,3\,4), (1\,3)\} \quad \text{and} \quad Y = \{(5\,7), (5\,6\,7)\}.$$
One can get even more specific since $\langle X \rangle \cong \mathcal{D}_4$ and $\langle Y \rangle \cong \mathcal{S}_3$. Therefore we have that
$$\langle (1\,2\,3\,4), (5\,7), (5\,6\,7), (1\,3) \rangle \cong \mathcal{D}_4 \times \mathcal{S}_3.$$
In particular, we get that this group has order 48.

EXERCISE 12.3. Prove that if G is an abelian group generated by $\{g_1, \ldots, g_n\}$ with
$$|G| = |g_1| \cdots |g_n|,$$
then
$$G = \langle g_1 \rangle \times \cdots \times \langle g_n \rangle .$$
SOLUTION

12.2 ACTIONS BY AUTOMORPHISMS

Semi-direct products take three bits of information, two groups G and H, and an action of G on H by automorphisms. We now define what the heck this last thing is.

DEFINITION 12.3. Let G and H be groups. Let $\bullet : G \circlearrowright H$ be an action. We say this action *is by automorphisms if* for all $g \in G$, $g \bullet$ is an automorphism (of H).

Another way to say this is to say that the permutation representation

$$\begin{aligned}\rho^{\bullet} &: G \to \mathcal{S}_H \\ &: g \mapsto g \bullet\end{aligned}$$

has image contained in the group of automorphism of H, $\mathrm{Aut}(H) \preccurlyeq \mathcal{S}_H$. So an action of a group G on a group H by automorphisms is the same thing as a homomorphism of groups

$$G \to \mathrm{Aut}(H).$$

In even more detail, a function $\bullet : G \circlearrowright H$ is an action by automorphisms iff (the multiplication in K is written in explicitly as $*$ for emphasis.)

(1) $\forall k \in K,\ e_G \bullet k = k$,

(2) $\forall k \in K, \forall g_1, g_2 \in G,\ (g_1 g_2) \bullet k = g_1 \bullet (g_2 \bullet k)$,

(3) $\forall k_1, k_2 \in K, \forall g \in G,\ g \bullet (k_1 * k_1) = (g \bullet k_1) * (g \bullet k_1)$.

Of course only the third requirement is new, and the first two are simply the usual requirements that $\bullet$ is an action in the usual sense. For such an action by automorphism, we have the following three properties which are simply translations of the three part of Theorem8.10 which can be used since $g \bullet$ is a homomorphism (actually an automorphism) of groups, $H \to H$

(1) $\forall g \in G,\ g \bullet e_K = e_K$

(2) $\forall g \in G, \forall h \in H\ g \bullet \left(h^{-1}\right) = \left(g \bullet h\right)^{-1}$

(3) $\forall g \in G, \forall h \in H, \forall n \in \mathbb{Z}\ g \bullet h^n = \left(g \bullet h\right)^n$.

When compared to the study of groups in general, the study of Abelian groups in and of themselves is pretty trivial. However, using our semi-direct product construction, we can form very exciting examples of groups simply from a pair of a finite abelian groups and a reasonable action. This is due, in part, to the fact that despite the fact that abelian groups themselves aren't all that exciting, abelian groups can have very interesting automorphism groups. For instance, the abelian groups $\mathbb{Z}_n \times \mathbb{Z}_n$ have pretty exciting automorphism groups, namely $\mathbf{GL}_2(\mathbb{Z}_n)$. For instance

$$\begin{aligned}&: \mathbb{Z}_6 \to \mathbf{GL}_2(\mathbb{Z}_5) \\ &: [1]_6 \mapsto \begin{pmatrix} [0]_5 & -[1]_5 \\ [1]_5 & [1]_5 \end{pmatrix}\end{aligned}$$

determines an action $\mathbb{Z}_6 \circlearrowright \mathbb{Z}_5^2$[1]. In this action we have, for instance, to do a few random calculations (leaving off the brackets around the guys in $\mathbb{Z}_5$ to save space, and not around $\mathbb{Z}_6$ so their distinction jumps out a bit more).

$$[1]_6 \bullet (1,0) = (0,1) \qquad [2]_6 \bullet (1,0) = [1]_6 \bullet (0,1) = (4,1)$$
$$[1]_6 \bullet (2,3) = (-3, 2+3) = (2,0) \qquad [2]_6 \bullet (2,3) = (0,2)$$

etc.

Some examples of (abstract) actions of one group on another by automorphisms include

(1) For any groups G and H the trivial action, defined by $g \bullet h := h$ is an action by automorphisms. It corresponds to the trivial homomorphism

$$G \to \mathrm{Aut}(H).$$

[1] There's nothing special about the integer 5 here; this works for any integer since the matrix written above always has order a divisor of 6 (and is exactly 6 except in the case that $n = 2$).

(2) Every group G conjugation, $g \bullet_C x := gxg^{-1}$ defines an action of G on itself by automorphisms.

(3) More generally, if $N \trianglelefteq G$ and $H \leqslant G$, then this conjugation

$$h \bullet_C x := hxh^{-1} \qquad (h \in H, x \in N)$$

defines an of H on N by automorphisms. We are about to see that, in a sense, this is the most general kind of an action by automorphisms. In more detail, for any action $\bullet \colon H \circlearrowright N$ by automorphisms, there is a group G, namely the semi–direct product of the groups H and N which is about to defined, so that the action $\bullet$ is equal to the action $\bullet_C$ defined here.

(4) For any group G, the automorphism group of G, $\mathrm{Aut}(G)$, acts on G by automorphisms:

$$\varphi \bullet g := \varphi(g).$$

Of course the permutation representation of this action is simply the identity, $\mathrm{Aut}(G) \to \mathrm{Aut}(G)$.

EXERCISE 12.4. Let $\bullet \colon G \circlearrowright H$ be an action by automorphisms.

(1) Let $T \subseteq H$. Prove that

$$G_T \stackrel{\text{def}}{=} \{g \in G : \forall h \in H, g \bullet h = h\}$$

is a subgroup of G which is normal in G for $T = H$.
Hint! We've already done this. Don't be fooled into doing this again just we are in a new situation.

(2) Let $S \subseteq G$. Prove that

$$H^S \stackrel{\text{def}}{=} \{h \in H : \forall g \in S, g \bullet h = h\}$$

is a subgroup of H. Prove via example that H^G needn't be a normal subgroup of H.
Hint! Of course to give an action you need to explicitly define G, H, and an action $G \circlearrowright H$ by automorphisms. I would use[II] a group with an element whose centralizer isn't normal. For instance $\mathcal{S}_3$ has such an element.

SOLUTION

EXERCISE 12.5. Suppose that we have an action $G \circlearrowright H$ by automorphisms. Suppose that $\overline{H} \leqslant H$ is invariant under the action. That is for all $g \in G$, $g \bullet \overline{H} = \overline{H}$.

(1) Prove that

$$g \star \left(h\overline{H}\right) := (g \bullet h)\overline{H}$$

is a well-defined action of G on the set of cosets $H/\overline{H}$.

(2) Prove that if $\overline{H} \trianglelefteq H$, then the action defined in the first part is by automorphisms. (This makes sense the assumption guarantees that $H/\overline{H}$ is a group, and like any group, you can ask about its automorphisms.)

SOLUTION

EXAMPLE 12.4. For ease, let's denote $\mathbf{GL}_3\left(\mathbb{Z}_p\right)$ as G. We use the previous exercise to show how to build a chain of subgroups of G which end in $\mathrm{UT}_3\left(\mathbb{Z}_p\right)$. We consider the usual action of $\mathbf{GL}_3\left(\mathbb{Z}_p\right) \circlearrowright \mathbb{Z}_p^3$. Here we not only make use of the group structure of $\mathbf{GL}_3\left(\mathbb{Z}_p\right)$ (under matrix multiplication) but also of $\mathbb{Z}_p^3$ (under addition). For one

thing, this action is by automorphisms of $\mathbb{Z}_p^3$. Under this action, the stabilizer of the vector $\mathbf{e}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$ consists of the matrices of the form

$$G_{\mathbf{e}_1} = \left\{ \begin{pmatrix} 1 & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{pmatrix} \in G : A_{ij} \in \mathbb{Z}_p \right\}.$$

Of course the action $G_{\mathbf{e}_1} \circlearrowright \mathbb{Z}_p^3$ therefore fixes the subgroup $\langle \mathbf{e}_1 \rangle \preccurlyeq \mathbb{Z}_p^3$, which like every subgroup of an abelian group, is normal. Therefore by the exercise we get a well defined action of $G_{\mathbf{e}_1} \circlearrowright \mathbb{Z}_p^3/\langle \mathbf{e}_1 \rangle$. Of course this quotient can be generated by $\mathbf{e}_2' := \mathbf{e}_2 + \langle \mathbf{e}_1 \rangle$ and $\mathbf{e}_3' := \mathbf{e}_3 + \langle \mathbf{e}_1 \rangle$. What's the stabilizer of $\mathbf{e}_2'$? Well it consists of matrices $\mathbf{A}$ in $G_{\mathbf{e}_1}$ with

$$\mathbf{A}(\mathbf{e}_2) + \langle \mathbf{e}_1 \rangle = \mathbf{A}(\mathbf{e}_2 + \langle \mathbf{e}_1 \rangle) = \mathbf{e}_2 + \langle \mathbf{e}_1 \rangle.$$

Using the definition of a coset (working in $\mathbb{Z}_p^3$, not $\mathbf{GL}_3\left(\mathbb{Z}_p\right)$), this means that $\mathbf{A}(\mathbf{e}_2) - \mathbf{e}_2$ must be a multiple of $\mathbf{e}_1$. Hence

$$(G_{\mathbf{e}_1})_{\mathbf{e}_2'} = \left\{ \begin{pmatrix} 1 & A_{12} & A_{13} \\ 0 & 1 & A_{23} \\ 0 & 0 & A_{33} \end{pmatrix} \in G : A_{ij} \in \mathbb{Z}_p \right\}.$$

Continuing in this way, by this same exercise we get an action of $(G_{\mathbf{e}_1})_{\mathbf{e}_2'}$ on

$$\left(\mathbb{Z}_p^3/\langle \mathbf{e}_1 \rangle\right)/\langle \mathbf{e}_2' \rangle = \mathbb{Z}_p^3/\langle \mathbf{e}_1, \mathbf{e}_2 \rangle$$

where the equality is due to the isomorphism theorems. This quotient is generated now only by the image of $\mathbf{e}_3$, denoted $\mathbf{e}_3''$. Now a similar argument as the last one shows that

$$((G_{\mathbf{e}_1})_{\mathbf{e}_2'})_{\mathbf{e}_3''} = \left\{ \begin{pmatrix} 1 & A_{12} & A_{13} \\ 0 & 1 & A_{23} \\ 0 & 0 & 1 \end{pmatrix} \in G : A_{ij} \in \mathbb{Z}_p \right\}$$

which is the group we have always called $\mathrm{UT}_3\left(\mathbb{Z}_p\right)$.

In more generality for n an arbitrary integer, we can continue this chain of subgroups of $\mathbf{GL}_n\left(\mathbb{Z}_p\right)$ but considering the stabilzer on smaller and smaller quotients of $\mathbb{Z}_p^n$. You should be convinced that the eventual stabilizer is always $\mathrm{UT}_n\left(\mathbb{Z}_p\right)$, which, as the stabilizer of a point under some action, gives an elegant proof that these sets are always subgroups of $\mathbf{GL}_n\left(\mathbb{Z}_p\right)$, and always having determinant 1, they are also also subgroups of $\mathbf{SL}_n\left(\mathbb{Z}_p\right)$ as well.

Continuing in the theme of the previous example we prove the following:[III]

Example /Theorem 12.5. The group $\mathrm{UT}_n\left(\mathbb{Z}_p\right)$ has order equal to $p^{\binom{n}{2}}$, where, as usual,

$$\binom{n}{2} = \frac{n \cdot (n-1)}{2}.$$

[III] This is also not 100% relevant to this chapter but it doesn't really fit in elsewhere.

Proof. One way to see that is due to the fact that $\binom{n}{2}$ can be thought of as the sum of the first $n-1$ numbers. Meanwhile we have an element of $\mathrm{UT}_n\left(\mathbb{Z}_p\right)$ can be any of the matrices of the form

$$\begin{pmatrix} 1 & * & * & \cdots & * & * \\ 0 & 1 & * & \cdots & * & * \\ 0 & 0 & 1 & \cdots & * & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & * \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

where the $*$'s are free to take on any of the p elements of $\mathbb{Z}_p$ they choose. Reading across the top row, there are $n-1$ $*$'s, the second there are $n-2$ $*$'s, etc. Thus there are a total of

$$\binom{n}{2} = (n-1) + (n-2) + \cdots + 2 + 1$$

$*$'s which each can take on p values. Therefore there are a total of $p^{\binom{n}{2}}$ elements of $\mathrm{UT}_3\left(\mathbb{Z}_p\right)$, as promised. ∎

In these exercises, λ denotes the Cayley action: $\lambda\colon G \to \mathcal{S}_G$ given by

$$\lambda\colon g \mapsto (h = g \cdot h).$$

(Or written using the more convenient subscript notation, $\lambda_g(h) = g \cdot h$.) Do not forget that λ is a monomorphism of groups (this is exactly the statement of Cayley's theorem).

Exercise 12.6. Suppose $G \leqslant \widehat{G}$. Define a monomorphism

$$\mathrm{N}_{\widehat{G}}(G)/C_{\widehat{G}}(G) \to \mathrm{Aut}\,(G)\,.$$

SOLUTION

Hint! Define a homomorphism from $\mathrm{N}_{\widehat{G}}(G) \to \mathrm{Aut}\,(G)$ whose kernel is $C_{\widehat{G}}(G)$.

Exercise 12.7. Prove that $\mathrm{Aut}\,(G)$ is isomorphic to a subgroup of $\mathcal{S}_G$ which normalizes $\lambda(G)$.
Hint! This may seem intimidating, but a direct check shows that

$$\varphi \circ \lambda_g \circ \varphi^{-1} = \lambda_{\varphi(g)} \tag{1}$$

which immediately implies the exercise (why?).

SOLUTION

This means that $\mathrm{Aut}\,(G)$ is (isomorphic to) a subgroup of $\mathrm{N}_{\mathcal{S}_G}(\lambda(G))$.[IV] Actually, a bit later, we'll see that $\mathrm{Aut}\,(G)$ consists precisely of the elements of this normalizer which fix $e_G \in G$.

[IV] Actually, since $\mathrm{Aut}\,(G)$ is literally a subset of the bijection of G, $\mathcal{S}_G$ it is an honest to goodness subgroup, not just isomorphic to one.

Exercise 12.8. Show that under the identification $\mathrm{Aut}(G) \preccurlyeq \mathrm{N}_{\mathcal{S}_G}(\lambda(G))$ from Exercise 12.7, that $\mathrm{Aut}(G)$ does not meet the centralizer $C_{\mathcal{S}_G}(\lambda(G))$.
Hint! Equation (1) implies that φ commutes with λ_g iff $\varphi(g) = g$.
SOLUTION

Exercise 12.9. Prove that the group

$$\mathrm{N}_{\mathcal{S}_G}(\lambda(G)) \big/ C_{\mathcal{S}_G}(\lambda(G)) \cong \mathrm{Aut}(G).$$

Hint! We have already proven that $\mathrm{Aut}(G)$ (which is isomorphic to $\mathrm{Aut}(\lambda(G))$ since $G \cong \lambda(G)$) is isomorphic to a subgroup of $\mathcal{S}_G$ which normalizes G. This gives us the inverse of the map from Exercise 12.6 (applied to the special case that $G = \lambda(G)$ and $\widehat{G} = \mathcal{S}_G$).
SOLUTION

This previous exercise, while not too useful for us humans,[V] is incredibly useful for computers. This is in part due to the fact that in Section 10.10 we've come up with an algorithm which computes stabilizers, which normalizers and (to an extent) centralizers are specific examples. This gives us an algorithm which calculates $\mathrm{Aut}(G)$ for any finite group G.

We further classify this normalizer in Example/Theorem 12.34 coming up on page 677.

12.3 Main Definitions of SDP's

As we already pointed out at the beginning of the previous section, a semi-direct product takes 3 bits of information:

- A group H
- A group K
- An action of $\bullet\colon K \circlearrowright H$ by automorphisms

and forms a new group

$$H \rtimes_\bullet K$$

which as a set is just $H \times K$[I] but with the multiplication skewed by the action:

$$(h,k) \cdot (h',k') := (h(k \bullet h'), kk').$$

Of course the first thing we need to do is prove that this is a group.

Theorem 12.6. *$H \rtimes_\bullet K$ is a group.*

Proof. Since $H \rtimes_\bullet K$ is not a subset of any known group, we cannot apply any of the subgroup tests and we have actually construct everything. Indeed the associative property, which is most of the time trivial to prove, become the biggest headache.

[V] Image trying to find the normalizer of a $\mathcal{S}_4$ (identified with) inside of $\mathcal{S}_{24}$.
[I] This is the first time where the underlying set of a group is not the explicit notation for a group. We will always reserve the symbol $H \times K$ to mean the usual direct product of groups.

(1) For notational ease, we explicitly write the operation in H as $*$ and we use the notation ${}^k h$ instead of $k \bullet h$ for this part alone! With these notations, we have that

$$ {}^k({}^{k'}h) = {}^{kk'}h \quad \text{(Associativity of an action)} \tag{2} $$

$$ {}^k(h * h') = ({}^k h) * ({}^k h') \qquad \text{(Action is by Auts).} \tag{3} $$

To see that the multiplication is associative we have

$$ \begin{aligned} \Big((h,k)\cdot(h',k')\Big)\cdot(h'',k'') &= \Big(h * {}^k h', kk'\Big)\cdot(h'',k'') \\ &= (h * {}^k h' * {}^{kk'}h'', kk'k'') \end{aligned} $$

and at the same time we have that

$$ \begin{aligned} (h,k)\cdot\Big((h',k')\cdot(h'',k'')\Big) &= (h,k)\cdot(h' * {}^{k'}h'', k'k'') \\ &= (h * {}^k(h' * {}^{k'}h''), kk'k''). \end{aligned} $$

Therefore, we simply need to show the equality of ${}^k(h' * {}^{k'}h'')$ and ${}^k h' * {}^{kk'}h''$, but these follow directly from the two points we started with.

(2) The identity is (e_G, e_H). Indeed we have that (since the identity acts as the identity)

$$ (e_H, e_K)\cdot(h,k) = (e_H(e_K \bullet h), e_K k) = (e_H h, k) = (h,k). $$

and (since the action is by automorphisms, we have $k \bullet e_H = e_H$)

$$ (h,k)\cdot(e_H, e_K) = (h(k \bullet e_H), k) = (h,k). $$

(3) Given an element $(h,k) \in H$, I claim that the inverse is $(k^{-1} \bullet h^{-1}, k^{-1})$. Indeed,

$$ \begin{aligned} (h,k)\cdot(k^{-1} \bullet h, k^{-1}) &= (h(k \bullet (k^{-1} \bullet h^{-1})), e_K) \\ &= (h(kk^{-1} \bullet h^{-1}), e_K) \\ &= (hh^{-1}, e_K) = (e_H, e_K). \end{aligned} $$

■

To help us get our foot in the door of intuition with these things note that we have

$$ (h, e_K)\cdot(h', e_K) = (hh', e_K) \qquad (e_H, k)(e_H, k') = (e_H, kk') $$

just like for the direct product of two groups.

Example 12.7. Define an action, $\bullet\colon \mathbb{Z}_2 \circlearrowright \mathbb{Z}_n$ by taking every element of $\mathbb{Z}_n$ to its negative, i.e.

$$ [\,1\,]_2 \bullet [\,a\,]_n = [\,-a\,]_n. $$

This action is by automorphisms since $\mathbb{Z}_n$ is an abelian group. Then it is easy to see that

$$ \mathcal{D}_n \cong \mathbb{Z}_n \rtimes_\bullet \mathbb{Z}_2 $$

Indeed, the isomorphism is given by

$$\begin{aligned}\varphi\colon \mathcal{D}_n &\to \mathbb{Z}_n \rtimes_\bullet \mathbb{Z}_2\\ \colon \rho^i r^\epsilon &\mapsto ([\,i\,]_n, [\,\epsilon\,]_2).\end{aligned}$$

This is a homomorphism since

$$\rho^i r^\epsilon \rho^j r^\delta = \begin{cases}\rho^{i-j} r^{\delta+\epsilon} \text{ if } \epsilon = 1\\ \rho^{i+j} r^\delta \text{ if } \epsilon = 0\end{cases}$$

$$|\,s\,|\,\varphi\mapsto \begin{cases}([\,i-j\,]_n, [\,\delta+\epsilon\,]_2) & \text{if } \epsilon = 1\\ ([\,i+j\,]_n, [\,\delta\,]_2) & \text{if } \epsilon = 0\end{cases}$$

while at the same time

$$\begin{aligned}\varphi(\rho^i r^\epsilon)\varphi(\rho^j r^\delta) &= ([\,i\,]_n, [\,\epsilon\,]_2)([\,j\,]_n, [\,\delta\,]_2)\\ &= \begin{cases}([\,i-j\,]_n, [\,\delta+\epsilon\,]_2) & \text{if } \epsilon = 1\\ ([\,i+j\,]_n, [\,\delta\,]_2) & \text{if } \epsilon = 0.\end{cases}\end{aligned}$$

Moreover φ is clearly a bijection: it has inverse

$$([\,a\,]_n, [\,\epsilon\,]_2) \mapsto \rho^a r^\epsilon.$$

Example 12.8. If the action of K on H is trivial (that is $k \bullet h = h$ for all $k \in K$ and $h \in H$) then $H \rtimes_\bullet K = H \times K$.

Exercise 12.10. Suppose that H is an abelian group. Prove that

$$\mathcal{C}(H \rtimes_\bullet K) = H^K \times (\mathcal{C}(K) \cap K_H)$$

where as above in Exercise 12.4,

$$H^K \stackrel{\text{def}}{=} \{h \in H : \forall k \in K, k \bullet h = h\} \quad \text{and} \quad K_H \stackrel{\text{def}}{=} \{k \in K : \forall h \in H, k \bullet h = h\}.$$

SOLUTION

Hint! Like most center proofs, take this in bits.

In particular, this says that when H and K are both abelian (so that $\mathcal{C}(K) = K$) the center of a semi-direct product is the direct product of things acted on trivially and things *acting* trivially.

Example 12.9. In particular, this exercise says that if either the action is faithful or K is centerless, then the center of the semi–direct product consists of just the set of global fixed points of H under the action of K. For instance, if we consider the action[II] $t \bullet [\,a\,]_8 = [\,3a\,]_8$ of $\mathbb{Z}_2$ (written multiplicatively) on $\mathbb{Z}_8$, we see that this fits into this

[II] This really is an action since $3^2 = 9 = 1 \bmod 8$.

category, and the global fixed points of this action in $\mathbb{Z}_8$ is

$$\begin{aligned}\{[a]_8 : [3a]_8 = [a]_8\} = \{[a]_8 : [2a]_8 = [0]_8\} \\ = \{[0]_8, [4]_8\}.\end{aligned}$$

Hence the center of the resulting semi–direct product is isomorphic to $\mathbb{Z}_2$ (since it's a group with two elements). On the other hand, for the (faithful) action[III] $t \bullet [a]_8 = [5a]_8$, the global fixed point set consists of

$$\{[a]_8 : [5a]_8 = [a]_8\} = \{[a]_8 : [4a]_8 = [0]_8\} = \{[0]_8, [2]_8, [4]_8, [6]_8\}.$$

which is isomorphic to $\mathbb{Z}_4$. Therefore the resulting semi–direct products have non-isomorphic centers and are therefore non-isomorphic.

PROPOSITION 12.10. *The functions*

$$\begin{aligned}H &\to H \rtimes_\bullet K \\ h &\mapsto (h, e_K)\end{aligned}$$

and

$$\begin{aligned}K &\to H \rtimes_\bullet K \\ k &\mapsto (e_H, k)\end{aligned}$$

are monomorphisms

PROOF. Same proof as for direct products. ∎

As such, we will typically identify H and K as subgroups of $H \rtimes_\bullet K$.

REMARK 12.11. From now on, when no confusion can come from it, we identity H with the subgroup

$$\{(h, e_K) \mid h \in h\} \leqslant H \rtimes_\bullet K$$

and K with the subgroup

$$\{(e_H, k) \mid h \in h\} \leqslant H \rtimes_\bullet K.$$

As such, since we have $(h, e_K)(e_H, k) = (h, k)$ we have that

$$H \rtimes_\bullet K = HK.$$

The following proposition shows that under this identification the action, $\bullet$ is identified with conjugation.

PROPOSITION 12.12.

(1) *We have* $khk^{-1} = k \bullet h$.

[III] This really is an action since $5^2 = 25 = 1 \bmod 8$.

(2) *H is normal in $H \rtimes_\bullet K$.*

(3) $H \cap K = \{e\}$.

Proof. The proof **(1)** is a direct calculation:

$$\begin{aligned} khk^{-1} &= (e_H, k)(h, e_K)(e_H, k^{-1}) \\ &= (e_H, k)(h, k^{-1}) \\ &= (k \bullet h, kk^{-1}) \\ &= (k \bullet h, e_K) \\ &= k \bullet h. \end{aligned}$$

The proof of **(2)** follows from the fact that the semi-direct product is generated by H and K and both normalize H. Lastly, the proof of **(3)** is also obvious since H is a set of first components while K is a set of second components. ∎

Remark 12.13. If K were also normal, the semi-direct product would be a usual direct product by the direct product recognition theorem. As a matter of fact, that is essentially the only difference between general semi-direct products and direct products as we now will see.

We also have a converse, which will allow us to show that a TON of groups are really semi-direct products. Notice that the only difference between this and the Direct product recognition theorem is that we do not require that K be normal in G.

Theorem 12.14 (Semi-Direct Product Recognition Theorem)**.** *Suppose $H, K \leqslant G$. Then G is isomorphic to a semi-direct product of H and K (with the conjugacy action of K on H) iff*

SDP1 $H \cap K = \{e_G\}$

SDP2 $\langle H \cup K \rangle = G$

SDP3 $H \trianglelefteq G$[IV]

Proof. We have already shown that if G is a semi-direct product that it has these properties. On the other hand, since H is normal in G we an action of K on H given by $k \bullet h = khk^{-1}$.

Define

$$\begin{aligned} \varphi\colon H \rtimes_\bullet K &\to G \\ (h, k) &\mapsto hk. \end{aligned}$$

I claim that φ is an isomorphism.

- φ is a homomorphism since we have

$$\begin{aligned} \varphi((h, k) \cdot (h', k')) &= \varphi(h(k \bullet h'), kk') \\ &= h(k \bullet h')kk' \\ &= h(kh'k^{-1})kk' \\ &= hkh'k' \\ &= \varphi(h, k)\varphi(h', k') \end{aligned}$$

[IV] As you can see, this is the only difference between this and the direct product recognition theorem, Theorem 8.77.

- The rest of the proof is exactly the same as it was for direct product recognition theorem. The function φ is one-to-one since

$$hk = \varphi(h,k) = e_G \Rightarrow h = k^{-1} \Rightarrow h \text{ and } k = e_G.$$

- The function φ is surjective since every element of G has the form hk. ■

EXAMPLE 12.15. The binary tetrahedral group, $B\left(\blacktriangleright^3\right)$, is a semi-direct product of $K = \mathbb{Z}_3$ and $H = Q_8$ since

SDP1 $H \cap K = \{1\}$ since their orders (8 and 3) are relatively prime.

SDP2 $\langle H \cup K \rangle = B\left(\blacktriangleright^3\right)$ since the least common multiple of H and K is 24.

SDP3 We have shown in a couple of places that Q_8 is normal in $B\left(\blacktriangleright^3\right)$.

EXAMPLE 12.16. $\mathcal{A}_4$ is a semi-direct product of normal V (the Klein 4 group) and $\mathbb{Z}_3$ (the $\mathbb{Z}_3$ part corresponds to any $\underline{3}$–cycle).

EXAMPLE 12.17. $\mathcal{S}_4$ is a semi-direct product of normal V (the Klein 4 group) and $\mathcal{S}_3$ with the $\mathcal{S}_3$ part corresponding to the usual $\mathcal{S}_3$ sitting inside of $\mathcal{S}_4$.

EXAMPLE 12.18. $\mathcal{S}_n$ is a semi-direct product of normal $\mathcal{A}_n$ and $\mathbb{Z}_2 = \langle (1\,2) \rangle$.

EXAMPLE 12.19. Q_8 is not the semi-direct product of any two PNT subgroups. Indeed, every non-trivial subgroup of Q_8 contains -1, so there are no possible subgroups of Q_8 which would allow for **SDP1** to hold.

EXAMPLE 12.20. For $n > 4$, $\mathcal{A}_n$ is not the semi-direct product of any two PNT subgroups since in these situations $\mathcal{A}_n$ is simple and therefore contains no PNT, normal subgroups.

NOTATION 12.21. Since $\mathrm{Aut}(\mathbb{Z}_n^r) \cong \mathbf{GL}_r(\mathbb{Z}_n)$ via the isomorphism $\varphi \mapsto (\varphi(\mathbf{e}_1)|\cdots|\varphi(\mathbf{e}_n))$, we can express a semi–direct product of the form and action of $\mathbb{Z}_m \circlearrowleft \mathbb{Z}_n^r$ by automorphisms simply as a matrix $\mathbf{A} \in \mathbf{GL}_r(\mathbb{Z}_n)$, with $\mathbf{A}^m = \mathbf{I}_n$. Therefore the notation

$$\mathbb{Z}_n^r \rtimes_{\mathbf{A}} \mathbb{Z}_m$$

makes sense. If we write the (standard basis) generators of $\mathbb{Z}_n^r$ as usual as, $\mathbf{e}_1, \ldots, \mathbf{e}_r$ (written additively) and the generator of $\mathbb{Z}_m$ as y (written multiplicatively), then note that we have

$$y\,\mathbf{e}_i\,y^{-1} = A_{1i}\mathbf{e}_1 + \cdots + A_{ri}\mathbf{e}_r.$$

or if we choose to write things multiplicatively:

$$y\,\mathbf{e}_i\,y^{-1} = \mathbf{e}_1^{A_{1i}} \cdot \cdots \cdot \mathbf{e}_r^{A_{ri}}.$$

EXAMPLE 12.22. In this extended example, we play with a concrete example of a semi–direct product. I do not attack this problem in a systematic way, but simply attack it in the way that I would normally personally attack on my own. I keep trying to compute products until my intuition begins to build about this group.

- Consider

$$G := (\mathbb{Z}_3 \times \mathbb{Z}_3) \rtimes_{\left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)} \mathbb{Z}_4.$$

We denote the generators of $(\mathbb{Z}_3 \times \mathbb{Z}_3)$ as a and b and the generator of $\mathbb{Z}_4$ as t. We break with our tradition and write all generators multiplicatively. This means that $tat^{-1} = a^2b$ and $tbt^{-1} = ab$. Note that this works since the order of this matrix is 4 (the trace is $3 \equiv_3 = 0$).

Of course $|G| = 36$ (since, as a set, it's $\mathbb{Z}_3^2 \times \mathbb{Z}_4$). Also note that $t^2at^{-2} = a^{-1}$ and $t^2bt^{-2} = b^{-1}$. To get our hands dirty, let's compute the order of the element $at \in G$. (In all calculations I make implicit use of the fact that a and b commute and each have order 3.) Well,

$$\begin{aligned}(at)^2 &= atat \\ &= ata(t^{-1}t)t \\ &= atat^{-1}t^2 \\ &= a(a^2b)t^2 \\ &= a^3bt^2 \\ &= bt^2.\end{aligned}$$

and

$$\begin{aligned}\left(bt^2\right)^2 &= bt^2bt^2 \\ &= bt^2bt^{-2} \\ &= bb^{-1} \\ &= e_G.\end{aligned}$$

Therefore at has order 4. Since we have just shown that at has order 4, which is the same as the order of the element t, perhaps at and t are conjugate? (Of course if we would have realized this in the first place we wouldn't have had to have done the calculation we just did!) Let's play! (yayyy!) The most obvious things to check are ata^{-1} and btb^{-1}. Indeed

$$\begin{aligned}ata^{-1} &= ata^{-1}t^{-1}t \\ &= a\left(tat^{-1}\right)^{-1}t \\ &= a(a^2b)^{-1}t \\ &= aa^{-2}b^{-1}t \\ &= a^{-1}b^{-1}t \\ &= a^2b^2t.\end{aligned}$$

and

$$\begin{aligned}a^2ta^{-2} &= a\left[a^2b^2t\right]a^{-1} \\ &= a^2b^2(ata^{-1}) \\ &= a^2b^2(a^2b^2t) \\ &= abt.\end{aligned}$$

meanwhile

$$\begin{aligned} btb^{-1} &= btb^{-1}t^{-1}t \\ &= b\left(tbt^{-1}\right)^{-1}t \\ &= b(ab)^{-1}t \\ &= bb^{-1}a^{-1}t \\ &= a^{-1}t \\ &= a^2t. \end{aligned}$$

and

$$\begin{aligned} b^2tb^2 &= b\left[a^2t\right]b^{-1} \\ &= a^2(btb^{-1}) \\ &= a^2a^2t \\ &= at. \end{aligned}$$

Aha! Moreover putting these together we can also see that t is conjugate to bt as well. Indeed,

$$\begin{aligned} (ba^2)t(ba^2)^{-1} &= b(a^2ta^{-2})b^{-1} \\ &= b(abt)b^{-1} \\ &= ab(btb^{-1}) \\ &= aba^2t \\ &= bt. \end{aligned}$$

This should make it clear that conjugacy class of t in G consists of the 9 elements

$$cc_G(t) = \{a^ib^jt : 0 \leq i, j < 3\}.$$

At the same time we can see that t^2 is conjugate to all 9 elements of the form

$$cc_G(t^2) = \{a^ib^jt^2 : 0 \leq i, j < 3\}$$

while (you guessed it) t^3 is conjugate to all 9 elements of the form

$$cc_G(t^3) = \{a^ib^jt^3 : 0 \leq i, j < 3\}.$$

Lastly, we can see that the conjugacy classes of (they consist of 4 elements)

$$cc_G(a) = \{a, a^2b, a^2, ab^2\} \quad \text{and} \quad cc_G(b) = \{b, ab, b^2, a^2b^2\}$$

and of course, as always,

$$cc_G(e_G) = \{e_G\}.$$

We have now found all conjugacy classes in G. This proves that every element of G has order 1, 2, 3, or 4 (note the lack of spread!) and the order table is

G	n	1	2	3	4
	$\mathbf{o}_n$	1	9	8	18

If we compare this to $\mathcal{S}_3 \times \mathcal{S}_3$'s, which is another group of order 36, that we came up with on 415, which was

$\mathcal{S}_3 \times \mathcal{S}_3$	n	1	2	3	6
	$\mathbf{o}_n$	1	15	8	12

We now see how far apart these two groups are from being isomorphic! (In fact, if were were asked simply just this perhaps the fastest way of seeing that these two groups are not isomorphic is by noting that $\mathcal{S}_3 \times \mathcal{S}_3$ has no element of order 4 while this group clearly does. In fact, the generator $\times$ itself has order 4)

There is one thing these two groups do have in common however, namely they prove have 8 elements of order 3. We will be able to explain this in the next chapter: in the language of that chapter, since both contain normal Sylow 3–subgroups which are isomorphic to $\mathbb{Z}_3 \times \mathbb{Z}_3$.

- Next, let's create a permutation representation of this group G that we've been dealing with. For starters, since we have

$$b^2tb^{-2} = at \Rightarrow b^2tb^{-2}t^{-1}$$

we have that a is a word in b and t and so we can cut down the generators we need to simply a and t. Next, we have that Let's first compute

$$\begin{aligned} a^2ta^2 &= a^2ta^2t^{-1}t \\ &= a^2(a^2b)^2t \\ &= a^2a^4b^2t \\ &= b^2t. \end{aligned}$$

Using this, we now see that

$$\begin{aligned} (a^2t)^2 &= a^2ta^2t \\ &= b^2t^2 \end{aligned}$$

and

$$\begin{aligned} (a^2t)^3 &= (a^2\,t)\,b^2\,t^2 \\ &= a^2\,t\,b^2\,t^{-1}t^3 \\ &= a^2(ab)^2t^3 \\ &= a^4b^2t^3 \\ &= ab^2t^3. \end{aligned}$$

Hence,

$$\left\langle btb^{-1} \right\rangle = \left\langle a^2t \right\rangle = \{e, a^2t, b^2t^2, ab^2t^3\}.$$

As such,

$$\langle\, t \,\rangle \cap \langle\, btb^{-1} \,\rangle = \{e\}$$

and $\langle\, t \,\rangle$ is malnormal. There are 9 cosets of $\langle\, t \,\rangle$ in G, they are

$$\begin{array}{lll} \langle\, t \,\rangle & a\langle\, t \,\rangle & a^2\langle\, t \,\rangle \\ b\langle\, t \,\rangle & ba\langle\, t \,\rangle & ba^2\langle\, t \,\rangle \\ b^2\langle\, t \,\rangle & b^2a\langle\, t \,\rangle & b^2a^2\langle\, t \,\rangle . \end{array}$$

Now a and b permute these cosets according to the Cayley action (leaving off the $\langle\, t \,\rangle$ in the notation below)

$$a \mapsto \big(1\ a\ a^2\big)\big(b\ ba\ ba^2\big)\big(b^2\ b^2a\ b^2a^2\big) \quad \text{and} \quad b \mapsto \big(1\ b\ b^2\big)\big(a\ ba\ b^2a\big)\big(a^2\ ba^2\ b^2a^2\big).$$

(Actually, it's clear that t and a ⟨or b if you prefer it to a ⟩ generate G, so computing the image of b was a waste of time.) Meanwhile, for any $w \in \langle\, a, b \,\rangle$, we have

$$tw\langle\, t \,\rangle = twt^{-1}t\langle\, t \,\rangle = twt^{-1}(t\langle\, t \,\rangle) = twt^{-1}\langle\, t \,\rangle$$

and so finding the action of t on the cosets amounts to dealing with t's matrix:

$$\mathbf{A} := \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Indeed, since

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^2 = -\mathbf{I}_2 \quad \text{and} \quad \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^3 = -\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$$

we can, looking at the first columns, just read off the orbits of a and b (remembering to change to multiplicative notation):

$$t \mapsto \big(1\ \big)\big(a\ a^2b\ a^2\ ab^2\big)\big(b\ ab\ b^2\ a^2b^2\big).$$

Using the (implied) dictionary to make this seem a bit firmer:

$$a \mapsto (1\,2\,3)(4\,5\,6)(7\,8\,9) \qquad b \mapsto (1\,4\,7)(2\,5\,8)(3\,6\,9) \quad \text{and} \quad c \mapsto (1)(2\,6\,3\,8)(4\,5\,7\,9).$$

- Continuing with this same example, note that in G the elements a and b are **not** conjugate to each other. Note however, we would have a hard time telling these two elements apart. They both have the same order and each are conjugate to the same number of elements. Therefore, a question could be

  is there an automorphism of G which takes a to b?

  (Of course, if automorphism exists, it cannot be inner). We will show that indeed there is one by sticking G into a larger group. Let

$$\Phi := \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

denote the golden matrix. Over $\mathbb{Z}_3$ this matrix has order 8 and so defines an action $\mathbb{Z}_8 \circlearrowright \mathbb{Z}_3^2$ by automorphisms. Moreover, since Φ^2 is the matrix we've been dealing with we see that

$$G = \mathbb{Z}_4 \rtimes_{\mathbf{A}} \mathbb{Z}_3^2 \leqslant \mathbb{Z}_8 \rtimes_{\Phi} \mathbb{Z}_3^2,$$

which having index 2, is normal. Therefore conjugation by Φ induces an automorphism of G, and under this automorphism b gets sent to a.

For matrices, $\mathbf{A}$ and $\mathbf{A}'$ in $\mathbf{GL}_r(\mathbb{Z}_n)$, here are couple of different ways of telling that the relevant semi-direct products are isomorphic, i.e.

$$\mathbb{Z}_n^r \rtimes_{\mathbf{A}} \mathbb{Z}_m \cong \mathbb{Z}_n^r \rtimes_{\mathbf{A}'} \mathbb{Z}_m.$$

Example /Theorem 12.23.

(1) If the matrices $\mathbf{A}$ and $\mathbf{A}'$ are conjugate to each other in $\mathbf{GL}_r(\mathbb{Z}_n)$, then their relevant semi-direct products are isomorphic.

(2) If matrices $\mathbf{A}$ and $\mathbf{A}'$ generate the same subgroup of $\mathbf{GL}_r(\mathbb{Z}_n)$, then their relevant semi-direct products are isomorphic.

Proof.

(1) Suppose the matrices $\mathbf{A}$ and $\mathbf{A}'$ are conjugate to each other via the matrix $\mathbf{B}$. This just means that

$$\mathbf{B} \cdot \mathbf{A} \cdot \mathbf{B}^{-1} = \mathbf{A}'.$$

Then define the map

$$\begin{aligned} \varphi \colon \mathbb{Z}_n^r \rtimes_{\mathbf{A}} \mathbb{Z}_m &\to \mathbb{Z}_n^r \rtimes_{\mathbf{A}'} \mathbb{Z}_m. \\ : (\mathbf{v}, a) &\mapsto (\mathbf{B}(\mathbf{v}), a). \end{aligned}$$

This map is operation preserving (denoting the operation in the first sdp as $\cdot$ and the second as $*$ for emphasis) since

$$\begin{aligned} \varphi((\mathbf{v}, a) \cdot (\mathbf{w}, b)) &= \varphi((\mathbf{v} + \mathbf{A}^a(\mathbf{w}), a + b)) \\ &\mapsto (\mathbf{B}(\mathbf{v}) + \mathbf{B}(\mathbf{A}^a(\mathbf{w})), a + b) \\ &= (\mathbf{B}(\mathbf{v}) + \mathbf{A}'^a(\mathbf{B}(\mathbf{w})), a + b) \\ &= (\mathbf{B}(\mathbf{v}), a) * (\mathbf{B}(\mathbf{w}), b) \\ &= \varphi((\mathbf{v}, a)) * \varphi((\mathbf{w}, b)). \end{aligned}$$

Moreover, the map φ is clearly a bijection since $(\mathbf{v}, a) \mapsto (\mathbf{B}^{-1}(\mathbf{v}), a)$ is an inverse.

(2) The assumption that $\langle \mathbf{A} \rangle = \langle \mathbf{A}' \rangle$ means that there exists an integer α which is relatively prime to m so that $[\alpha]_m$ acts on $\mathbb{Z}_n^r$ as $\mathbf{A}'$ (and $\mathbf{A}^\alpha = \mathbf{A}'$). This time we define the map

$$\begin{aligned} \varphi \colon \mathbb{Z}_n^r \rtimes_{\mathbf{A}'} \mathbb{Z}_m &\to \mathbb{Z}_n^r \rtimes_{\mathbf{A}} \mathbb{Z}_m. \\ : (\mathbf{v}, [a]_m) &\mapsto (\mathbf{v}, \alpha \cdot [a]_m) \end{aligned}$$

in the opposite direction. We check that φ is an isomorphism. For starters, φ is operation preserving since

$$\begin{aligned} \varphi((\mathbf{v}, [a]_m) * (\mathbf{w}, [b]_m)) &= \varphi((\mathbf{v} + \mathbf{A}'^a(\mathbf{w}), [a + b]_m)) \\ &= (\mathbf{v} + \mathbf{A}'^a(\mathbf{w}), \alpha\,[a + b]_m) \\ &= (\mathbf{v} + \mathbf{A}^{\alpha \cdot a}(\mathbf{w}), \alpha\,[a + b]_m) \\ &= (\mathbf{v}, \alpha\,[a]_m) \cdot (\mathbf{w}, \alpha\,[b]_m) \\ &= \varphi((\mathbf{v}, [a]_m)) \cdot \varphi((\mathbf{w}, [b]_m)). \end{aligned}$$

This is a bijection since we have that α is relatively prime to m, so there is a (mod m) multiplicative inverse, α^{-1}, of α. Then we have that the map

$$\begin{aligned} \varphi^{-1} \colon \mathbb{Z}_n^r \rtimes_{\mathbf{A}} \mathbb{Z}_m &\to \mathbb{Z}_n^r \rtimes_{\mathbf{A}'} \mathbb{Z}_m. \\ : (\mathbf{v}, [a]_m) &\mapsto (\mathbf{v}, \alpha^{-1} \cdot [a]_m) \end{aligned}$$

is an inverse to φ. ∎

As a particular case of this Example/Theorem 12.23 **(2)** (since for any element $\mathbf{A}$ of any group we always have $x\langle \mathbf{A} \rangle = \langle \mathbf{A}^{-1} \rangle$) we know that we always have

$$\mathbb{Z}_n^r \rtimes_{\mathbf{A}} \mathbb{Z}_m \cong \mathbb{Z}_n^r \rtimes_{\mathbf{A}^{-1}} \mathbb{Z}_m.$$

For instance, looking back to our extended example, Example 12.22,

$$(\mathbb{Z}_3 \times \mathbb{Z}_3) \rtimes_{\left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)} \mathbb{Z}_4 \cong (\mathbb{Z}_3 \times \mathbb{Z}_3) \rtimes_{\left(\begin{smallmatrix} 1 & 2 \\ 2 & 2 \end{smallmatrix}\right)} \mathbb{Z}_4.$$

What would have really made that example even easier would have been if we'd have used the first part, Example/Theorem 12.23 **(1)** and the fact that in $\mathbf{GL}_2(\mathbb{Z}_3)$ the matrices

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

are conjugate to each other. (The fact that they are conjugate is a consequence of Lemma G.11 found in the appendix. However, it is not too difficult to find a matrix which conjugates one matrix to the other directly.) Then we'd have

$$(\mathbb{Z}_3 \times \mathbb{Z}_3) \rtimes_{\left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)} \mathbb{Z}_4 \cong (\mathbb{Z}_3 \times \mathbb{Z}_3) \rtimes_{\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)} \mathbb{Z}_4.$$

In this latter group we'd instead have the relations

$$tat^{-1} = b^{-1} \text{ and } tbt^{-1} = a$$

which look a lot nicer and easier to deal with than what we actually dealt with:

$$tat^{-1} = a^2 \cdot b \text{ and } tbt^{-1} = a \cdot b.$$

EXAMPLE 12.24. We prove that the group

$$\left\{ \begin{pmatrix} [\,a\,]_{p^2} & [\,b\,]_{p^2} \\ [\,0\,]_{p^2} & [\,1\,]_{p^2} \end{pmatrix} : [\,a\,]_p = [\,1\,]_p \right\}$$

from Exercise 6.6 back on page 323 is isomorphic to the the semi–direct product

$$\mathbb{Z}_{p^2} \rtimes_{(p+1)} \mathbb{Z}_p.$$

We take the copy of $\mathbb{Z}_{p^2}$ to be the group generated by

$$\mathbf{A} := \begin{pmatrix} [\,1\,]_{p^2} & [\,1\,]_{p^2} \\ [\,0\,]_{p^2} & [\,1\,]_{p^2} \end{pmatrix}.$$

This really is isomorphic to this group since $\mathbf{A}$ has order p^2.

We take the copy of $\mathbb{Z}_p$ to be the group generated by

$$\mathbf{B} := \begin{pmatrix} [\,p+1\,]_{p^2} & [\,0\,]_{p^2} \\ [\,0\,]_{p^2} & [\,1\,]_{p^2} \end{pmatrix}.$$

The fact that $\mathbf{B}$ (and hence the group generated by $\mathbf{B}$) has order p is due to the fact that $(1+p) \neq 1 \bmod p$ and

$$(1+p)^p \equiv_{p^2} 1 + p \cdot p \equiv_{p^2} 1.$$

Of course our copies of $\mathbb{Z}_p$ and $\mathbb{Z}_{p^2}$ have only the identity matrix in common.

Next, note that

$$\begin{aligned}\mathbf{BAB}^{-1} &= \begin{pmatrix}[\,p+1\,]_{p^2} & [\,0\,]_{p^2}\\ [\,0\,]_{p^2} & [\,1\,]_{p^2}\end{pmatrix}\begin{pmatrix}[\,1\,]_{p^2} & [\,1\,]_{p^2}\\ [\,0\,]_{p^2} & [\,1\,]_{p^2}\end{pmatrix}\begin{pmatrix}[\,p+1\,]_{p^2} & [\,0\,]_{p^2}\\ [\,0\,]_{p^2} & [\,1\,]_{p^2}\end{pmatrix}^{-1}\\ &= \begin{pmatrix}[\,1\,]_{p^2} & [\,p+1\,]_{p^2}\\ [\,0\,]_{p^2} & [\,1\,]_{p^2}\end{pmatrix}\\ &= \mathbf{A}^{p+1}.\end{aligned}$$

This completes the proof.

Exercise 12.11. Let G be the group

$$\left(\mathbb{Z}_p \times \mathbb{Z}_p\right) \rtimes_{\mathfrak{R}} \mathbb{Z}_p$$

where, as usual, $\mathfrak{R}$ is the matrix

$$\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$$

and p is some prime. Prove that $G \cong \mathrm{UT}_3\left(\mathbb{Z}_p\right)$.

SOLUTION

Exercise 12.12. Prove that $[\,p+1\,]_{p^2}$ has multiplicative order p in $\mathcal{U}(\mathbb{Z}_{p^2})$.

SOLUTION

12.4 (Application) Classification of Groups of Order p^3

In 10.8.2 we classified groups of order 8. In this section we classify groups of order p for odd primes, p.

Remark 12.25. It will turn out that in this case there are also five groups of order p^3, 2 non–abelian, 3 abelian, but the proof will be quite a bit different. In fact, for p and odd prime, every non–abelian group of order p^3 is a semi–direct product, a fact which is false for $p = 2$ due to the presence of the quaternions, Q_8.

Let G be a non-abelian group of order p^3. From Exercise 11.38 back on page 644, we saw that every the center of G and the commutator of G coincide and are each subgroups of G of order p. This may feel useless at first, but we are about to see this is quite a useful thing. Indeed, note that since $[G, G] \subseteq \mathcal{C}(G)$ we have that the element

$$y^{-1}x^{-1}yx := z_{x,y} \tag{4}$$

which as a commutator of elements of G, is a central element. Note that the element $z_{x,y}$ depends on x and y, but to ease up on notation we denote it simply as z.

We have

$$\begin{aligned} zxy &= xyz \\ &= xy(y^{-1}x^{-1}yx) \\ &= yx. \end{aligned}$$

That is $yxy^{-1} = zx$. Also,

$$\begin{aligned} y^2xy^{-2} &= yyxy^{-1}y^{-1} \\ &= y(zx)y^{-1} \\ &= zyxy^{-1} \qquad (z \text{ is central}) \\ &= zzx \\ &= z^2x. \end{aligned}$$

This implies, by multiplying both sides of the resulting equation on the right by y^2, that $y^2x = z^2xy^2$. Inductively, the same argument shows us that for every integer n,

$$y^nx = z^nxy^n. \tag{5}$$

As always, in the lemma below

$$\binom{n}{2} = \frac{n(n-1)}{2}.$$

Note that we can think of $\binom{n}{2}$ as the sum of the first $n-1$ integers. With this perspective, it's obvious without a calculation that

$$\binom{n}{2} + n = \binom{n+1}{2} \tag{6}$$

since both sides represent the sum of the first n integers.

LEMMA 12.26. *For any integer n,*

$$(xy)^n = z^{\binom{n}{2}}x^ny^n.$$

PROOF. We prove this by induction with base case $n = 2$. In this case

$$\begin{aligned} (xy)^2 &= xyxy \\ &= x(zxy)y \quad (yx = zxy) \\ &= xzxyy \\ &= zxxyy \quad (z \text{ is central}) \\ &= zx^2y^2. \end{aligned}$$

Inductively, assuming the result is true for n, then

$$\begin{aligned}
(xy)^{n+1} &= (xy)^n\, xy \\
&= z^{\binom{n}{2}} x^n y^n xy && \text{(Ind hyp)} \\
&= z^{\binom{n}{2}} x^n z^n xy^n y && \text{(Equation (5))} \\
&= z^{\binom{n}{2}} z^n x^n xy^n y && (z \text{ is central}) \\
&= z^{\binom{n}{2}+n} x^{n+1} y^{n+1} && \text{(Grouping exponents)} \\
&= z^{\binom{n+1}{2}} x^{n+1} y^{n+1} && \text{(Equation (6))}
\end{aligned}$$

which completes the induction. ∎

Next we note that

$$\binom{p}{2} = \frac{p\,(p-1)}{2}$$

which means that since we are assuming p is odd (this is the first use of the fact that p is odd!) we know that $\frac{p-1}{2}$ is an integer and so p divides $\binom{p}{2}$. Now since $\mathcal{C}(G)$ has order p, we have $z^{\binom{p}{2}} = e_G$, which implies that the pesky $z's$ which keep showing up in our equations disappear when we take things to the p^{th} power and we have $(xy)^p = x^p\, y^p$. Let's denote the function $G \to G$ given by $x \mapsto x^p$ as F_p. We have proven

Lemma 12.27. *In our current situation, F_p is a homomorphism.*

Remark 12.28. All of this certainly generalizes! It gives us some indication why groups with

$$[G, G] \subseteq \mathcal{C}(G)$$

are pretty sweet.

What is the image of F_p? Well, since G is non–abelian, we cannot have that $G/\mathcal{C}(G)$ is cyclic. Moreover, since every group of order p^2, which this quotient is an example of, is either cyclic or isomorphic to $\mathbb{Z}_p \times \mathbb{Z}_p$, $G/\mathcal{C}(G)$ must be isomorphic to the latter. But this means that for every $g \in G$ that $g^p \in \mathcal{C}(G)$. Therefore we have proven yet another lemma:

Lemma 12.29. *The homomorphism F_p from the last lemma has image contained in the center of G.*

We now go onto classify non–abelian groups based on whether or not F_p is trivial or not. There are two cases: either F_p is trivial, in which case every element of G has order p, or F_p is non-trivial, in which case G contains an element of order p^2.

(1) (**Every element of G has order p**) Let's denote z as a generator of $\mathcal{C}(G)$ and let x be a non-central element of G. Then since $|x| = 3$, and since x and z commute, $\langle x, y \rangle \cong \mathbb{Z}_p^2$. Moreover, since this is an index p subgroup of G, it must be normal in G. In particular, since G is assumed to have order p^3, $\langle x, y \rangle$ does not contain some element $y \in G$. Since every element of G is assumed to have order p, y has order p. In particular,

$$\langle x \rangle \cap \langle x, y \rangle = \{e\} \quad \text{and} \quad \langle y \rangle \cdot \langle x, y \rangle = G.$$

After replacing z with $z_{x,y}$ from (4), we can assume that we have $yxy^{-1} = zx$.

These facts prove that

PROPOSITION 12.30. *Every group in this first case is isomorphic to* $\mathrm{UT}_3\left(\mathbb{Z}_p\right)$.

PROOF. All of our facts show that such a group is a semi–direct product of the form from Exercise 12.11. ■

(2) (**Some element of** G **has order** p^2) Let $x \in G$ be an element of order p^2. Since F_p has image contained in the center, we have $x^p \in \mathcal{C}(G)$. Moreover since $\langle x \rangle$ has index p, it is a normal subgroup of G. Moreover all $p-1$ elements of $\langle x \rangle$ whose order is p are central.

Next, since F_p in this case is non-trivial, it surjects onto the center (a group of order p). This means, by KIT, that this map has p^2 elements in its kernel, i.e. there are p^2-1 elements of G of order p. In particular, there is some element, besides the $p-1$ such elements in the center, $y \in G$ whose order is p. Therefore, the set of non-identity elements of $\langle y \rangle$, (all of which have order p and could function as a generator of $\langle y \rangle \cong \mathbb{Z}_p$) has empty intersection with $\mathcal{C}(G)$. But since all elements of order p in $\langle x \rangle$ **are** central, we have that $\langle x \rangle \cap \langle y \rangle = \{e_G\}$.

Choosing y appropriately, we can assume that $z_{x,y} = x^p$. That is we are assuming that $yxy^{-1} = x^p x = x^{p+1}$. All of these facts guarantee us that

PROPOSITION 12.31. *Every group in this second case is isomorphic to* Af_p.

PROOF. All of our facts show that such a group is a semi–direct product of the form from Example 12.24. ■

We have therefore finished our classification of groups of order p^3 for p and odd prime:

Every abelian group of order p^3 is isomorphic to exactly one of the following three groups:

(a) $\mathbb{Z}_{p^3}$
(b) $\mathbb{Z}_{p^2} \times \mathbb{Z}_p$
(c) $\mathbb{Z}_p \times \mathbb{Z}_p \times \mathbb{Z}_p$

Every non–abelian group of order p^3 is isomorphic to exactly one of the following two groups:

(a) $\mathrm{UT}_3\left(\mathbb{Z}_p\right)$
(b) Af_p

12.5 MORE EXAMPLES

12.5.1 THE TWO MISSING GROUPS OF ORDER 24.

In Section 8.5.1 we discussed groups of order 24. In addition to not showing that the groups presented in the table in that section, Table 8.2 back on page 418, were the only groups of order 24 (up to isomorphism) which is something we will not do, we never even defined the last two groups which (there) were simply denoted G_{24}^1 and G_{24}^2. It turns out that these two groups are semi-direct products of smaller groups.[I]

We repeat this table for convenience

[I] Nowhere in the text do we actually prove that these are the only groups of order 24!

G	$\mathcal{S}_4$	$B\left(\blacktriangleright^3\right)$	$B\mathcal{D}_6$	$\mathcal{D}_{12}$	$\mathcal{D}_6^{\times 2}$	$\mathcal{Q}_8^{\times 3}$	$\mathcal{D}_4^{\times 3}$	$\mathcal{A}_4^{\times 2}$	$B\mathcal{D}_3^{\times 2}$	$\mathcal{S}_3^{\times 4}$	G_{24}^1	G_{24}^2
LPO(G)	4	6	12	12	6	12	12	6	6	12	12	6
$\mathcal{C}(G)$	$\mathfrak{e}$	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}_2$	$\mathbb{Z}_2^2$	$\mathbb{Z}_6$	$\mathbb{Z}_6$	$\mathbb{Z}_2$	$\mathbb{Z}_2^2$	$\mathbb{Z}_4$	$\mathbb{Z}_4$	$\mathbb{Z}_2$
$\mathbf{o}_2(G)$	9	1	1	13	15	1	5	7	3	7	1	9

Table 12.1: All Non–Abelian Groups of Order 24=Table 8.2.

12.12.5.1.0 G_{24}^1

The first group listed is a semi-direct product

$$\mathbb{Z}_3 \rtimes \mathbb{Z}_8$$

where the action is taken to be the (one and only) non-trivial action of $\mathbb{Z}_8$ on $\mathbb{Z}_3$, $\mathbb{Z}_8 \circlearrowright \mathbb{Z}_3$, by automorphisms:

$$[\,a\,]_8 \bullet [\,b\,]_3 = (-1)^a \cdot [\,b\,]_3 .$$

This clearly is an action by automorphisms.

As a semi-direct product of a group of order 8 and a group of order 3 this dude really has order 24. Let's verify the other three data points about this group from the table, Table 12.1, i.e. that the group has a center isomorphic to $\mathbb{Z}_4$, it has an element of order 12, and that it contains a unique element of order 2. To verify these facts I may as well draw a conjugation diagram. I switch to using multiplicative notation while denoting the (a) generator of the $\mathbb{Z}_3$ factor as a and of the $\mathbb{Z}_8$ factor by t. We use the fact that

$$\begin{aligned} ata^{-1} &= ata^{-1}t^{-1}t \\ &= a\left(ta^{-1}t^{-1}\right)t \\ &= a\left(tat^{-1}\right)^{-1}t \\ &= a\left(a^{-1}\right)^{-1}t \\ &= aat \\ &= a^2t \end{aligned}$$

repeatedly.

$\mathbf{G}_{24}^1$	n	1	2	3	4	6	8	12
	$\mathbf{o}_n$	1	1	2	2	2	12	4

Table 12.2: Order table is read off from conjugation diagram, Figure 12.1. Notice how many more elements of order 8 and 12 there are then there are of 1, 2, 3, and 4! This indicates that many distinct elements of this group have common powers.

- Nothing in the acted on group is globally fixed, but all of $2\mathbb{Z}_8 \cong \mathbb{Z}_4$ in the acting group are in the kernel of the action. Therefore, by Exercise 12.10, the center of G_{24}^1 is isomorphic to $\mathbb{Z}_4$ as promised.

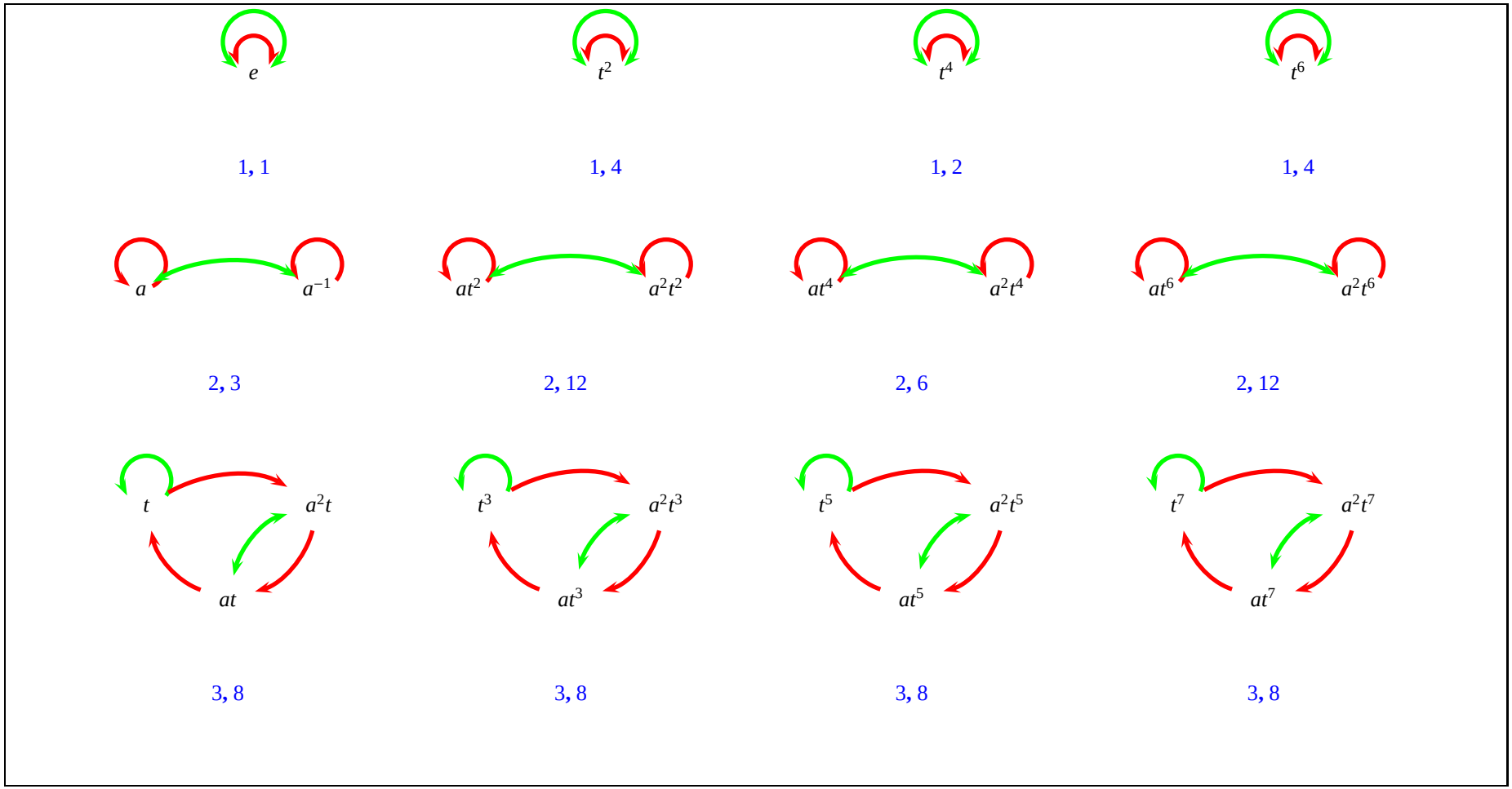


Figure 12.1: Conjugation by a corresponds to the red lines and conjugation by t corresponds to the green lines. Since t^2 is central we do not need to put arrow heads on the green lines. Notice that this conjugation table is not as useful as others. This is mostly due to the fact that the center is fairly sizeable and so that conjugacy classes themselves end up being rather small.

- Since G_{24}^1 is not abelian it cannot be cyclic, so the largest possible order an element of this group can have is 12. Let's verify that this group indeed has an element of order 12. This isn't hard to find since t^2 and a commute and have orders 4 and 3 respectively, we know that their product, at^2, has order 12.

- Lastly we verify that G_{24}^1 contains 1 and only 1 element of order 2. Of course t^4 is an element of order 2, so that much is done, but why is this the only element of order 2 in this group? Of course we could do this by brute force, for instance in Figure 12.1 we draw out the entire conjugation diagram and then we use this to record the order table, Table 12.2. However there is certainly a faster way to pull out this one piece of information. Indeed, let $w \in \mathbb{Z}_3$ and $s \in \mathbb{Z}_8$ we so that (w, s) has order

2 in G_{24}^1. Then

$$(w,s)^2 = (w,s)(w,s) \tag{7}$$
$$= (w \cdot (s \bullet w), s^2) \tag{8}$$
$$\Rightarrow w \cdot (s \bullet w) = e_{\mathbb{Z}_3} \quad \text{and} \quad s^2 = e_{\mathbb{Z}_8}. \tag{9}$$

But t^4 is the unique element of $\mathbb{Z}_8$ which has order 2, therefore $s = t^4$ or $s = e_{\mathbb{Z}_8}$. Since both of these elements are central in G_{24}^1 we have that $(s \bullet w) = w$ and so

$$w \cdot (s \bullet w) = e_{\mathbb{Z}_3} \Rightarrow w^2 = e_{\mathbb{Z}_3}.$$

But the only way this is possible is if $w = e_{\mathbb{Z}_3}$. Therefore, since we want (w,s) to have order 2, we must have $s \neq e_{\mathbb{Z}_8}$ since otherwise (w,s) would be the identity of the bigger group G_{24}^1 and hence not have order 2. Therefore we have that $(w,s) = (e_{\mathbb{Z}_8}, t^4)$ like I promised, and there is one and only one element of order 2 in G_{24}^1.

12.12.5.1.0 G_{24}^2

$$\mathbb{Z}_3 \rtimes \mathcal{D}_4$$

where the action is

$$\rho \bullet a = a^{-1} \quad \text{and} \quad \mathbf{r} \bullet a = a.$$

More detail to come

12.5.2 Holomorph of a Group

Obviously the largest group which can act on a given group G faithfully, by automorphisms, is the entire automorphism group, $\mathrm{Aut}(G)$. In this section, it is always implied that $\mathrm{Aut}(G)$ acts in this way: $\varphi \bullet g := \varphi(g)$.

Definition 12.32. The Holomorph of a group G, denoted $\mathrm{Hol}(G)$, is by definition

$$\mathrm{Hol}(G) \stackrel{\text{def}}{=} G \rtimes \mathrm{Aut}(G).$$

For us, this will only be interesting in the simplest cases of elementary abelian p–groups: $G = \mathbb{Z}_p^n$, where p is a prime. For instance, in the case that $n = 1$, since $\mathrm{Aut}\left(\mathbb{Z}_p\right) \cong \mathcal{U}(\mathbb{Z}_p)$, the groups $\mathrm{Hol}\left(\mathbb{Z}_p\right)$ have $(p-1)p$ elements. For instance, it's easy to see that $\mathrm{Hol}(\mathbb{Z}_3) \cong \mathbb{Z}_3$.

Example /Theorem 12.33. The map

$$\varphi : \mathrm{Hol}\left(\mathbb{Z}_p^n\right) \to \mathbf{GL}_{n+1}\left(\mathbb{Z}_p\right)$$
$$: (\mathbf{A}, \mathbf{v}) \mapsto \begin{pmatrix} \mathbf{A} & \mathbf{v} \\ \mathbf{0} & 1 \end{pmatrix}$$

defines an isomorphism from $\mathrm{Hol}\left(\mathbb{Z}_p^n\right)$ onto the set of matrices in $\mathbf{GL}_{n+1}\left(\mathbb{Z}_p\right)$ whose $n+1$ row vector is $\mathbf{e}_{n+1}$.

Proof. {Supply} ∎

For instance, in the case that $n = 1$ we have an isomorphism from $\text{Hol}\left(\mathbb{Z}_p\right)$ onto matrices of the form

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$$

in $\mathbf{GL}_2\left(\mathbb{Z}_p\right)$.

And for $n = 2$, we have an isomorphism from $\text{Hol}\left(\mathbb{Z}_p^2\right)$ onto matrices of the form

$$\begin{pmatrix} a & b & v_1 \\ c & d & v_2 \\ 0 & 0 & 1 \end{pmatrix}$$

in $\mathbf{GL}_3\left(\mathbb{Z}_p\right)$. For instance, we've seen a few times already that the set of matrices which fix $\mathbf{e}_1$ in $\mathbf{SL}_3\left(\mathbb{Z}_2\right)$ are isomorphic to $\mathcal{S}_4$. This group is clearly isomorphic to the above. Therefore, we have

$$\text{Hol}\left(\mathbb{Z}_2^2\right) \cong \mathcal{S}_4.$$

Another description of $\text{Hol}\left(\mathbb{Z}_p\right)$ is as a permutation group. Recall, for a prime p, from Theorem 5.75 back on page 289 $\mathcal{U}(\mathbb{Z}_p)$ is a cyclic group. As such, there exists a generator $x \in \mathcal{U}(\mathbb{Z}_p)$ (i,e, an element of multiplicative order $p-1$). For notational ease, we think of x as an honest to goodness integer,[II] but we remember to take all of ours results mod p. Then we have that

$$\text{Hol}\left(\mathbb{Z}_p\right) \cong \left\langle \left(1\ 2\ 3 \cdots\ p\right), \left(1\ x\ x^2\ x^3 \cdots x^{p-2}\right) \right\rangle.$$

For instance, for $p = 7$, 3 is a generator and following it around we have

$$3^0 = \textcircled{1} \bmod 7,\ 3^1 = \textcircled{3} \bmod 7,\ 3^2 = \textcircled{2} \bmod 7,\ 3^3 = \textcircled{6} \bmod 7,\ 3^4 = \textcircled{4} \bmod 7,\ 3^5 = \textcircled{5} \bmod 7.$$

As such we have that

$$\text{Hol}\left(\mathbb{Z}_7\right) \cong \langle (1\ 2\ 3\ 4\ 5\ 6\ 7), (1\ 3\ 2\ 6\ 4\ 5) \rangle.$$

Indeed, we have for any group G of finite order n,

$$\text{Hol}\left(G\right) \preccurlyeq \mathcal{S}_n$$

via Exercise 12.7 back on page 657, and its easy to see from this that quite often this is computed as the normalizer of G in $\mathcal{S}_n$.

Again, we denote the Cayley representation of a group G as

$$\lambda \colon G \to \mathcal{S}_G.$$

That is $\lambda_g(h) = gh$.

Example /Theorem 12.34.

$$\text{Hol}\left(G\right) \cong \text{N}_{\mathcal{S}_G}(\lambda(G)).$$

[II] To be technical we are using the dictionary $\mathbb{Z}_p \mid s \mid \mod p \to \mathbb{Z}$.

Remark 12.35. Like many example/theorems, I'm proving and dealing with this topic not so much for its usefulness, but because of the juxtaposition of abstraction and left brained thinking: it involves four groups at once: $\mathcal{S}_G$, $\mathrm{Aut}(G)$, G, and $\mathrm{Hol}(G)$ as well as the concept of the semi-direct products, but despite all this the proof amounts to a follow your nose left brained proof! Again, try not to be intimidated, but gather the relevant definitions and "follow your nose."

Proof. We already know that there are copies of the automorphism group of G, $\mathrm{Aut}(G)$, and G inside of $\mathcal{S}_G$.[III] Moreover since every automorphism fixes the identity of G and only the identity of G has $\lambda_g = \mathbf{1}_G$ we have that these two copies have trivial intersection. Next, we've also proven that for any automorphism φ of G that we have

$$\varphi \circ \lambda_a \circ \varphi^{-1} = \lambda_{\varphi(a)} \in \lambda(G)$$

implying that $\mathrm{Aut}(G)$ normalizes $\lambda(G)$. Therefore $\mathrm{Hol}(G) \leqslant \mathcal{S}_n$. Moreover, denoting $\mathrm{N}_{\mathcal{S}_G}(\lambda(G))$ as N, we have also proven $\mathrm{Hol}(G) \leqslant N$.

Now assume that $\alpha \in \mathcal{S}_G$ normalizes $\lambda(G)$. We want to prove that

$$\exists g \in G, \varphi \in \mathrm{Aut}(G) \text{ s.t. } \alpha = \lambda_g \circ \varphi. \tag{10}$$

If we consider $\lambda_{\alpha(e_G)^{-1}} \circ \alpha$, we see that it takes the identity of G, e_G back to e_G since

$$\lambda_{\alpha(e_G)^{-1}}(\alpha(e_G)) = \alpha(e_G)^{-1} \cdot \alpha(e_G) = e_G.$$

Moreover, as the composition of elements of I now claim that $\alpha' := \lambda_{\alpha(e_G)^{-1}} \circ \alpha$ is an automorphism of G (we at least have a shot since we proved that it fixes the identity of G).

Since α' normalizes $\lambda(G)$ we know that for every $g \in G$ there exists a $\overline{g} \in G$ (depending on g) so that

$$\alpha' \circ \lambda_g \circ \alpha'^{-1} = \lambda_{\overline{g}}. \tag{11}$$

I claim that for all $g \in G$, $\overline{g} = \varphi(g)$. Indeed since (11) is an equality of functions, we can apply both sides to our favorite element of their domain, which is G, e_G to get an equality in the group G.

$$\begin{aligned}
\alpha'(g) &= \alpha'(\lambda_g(e_G)) \\
&= \alpha'(\lambda_g(\alpha'^{-1}(e_G)) && \text{(Since } \alpha' \text{ fixes } e_G\text{)} \\
&= \lambda_{\overline{g}}(e_G) && \text{(This is o.s. of (??))} \\
&= \overline{g} \cdot e_G && \text{(Def'n of } \lambda\text{)} \\
&= \overline{g}.
\end{aligned}$$

Therefore my claim is true.

Lastly, we also have for all $a, b \in G$

$$\begin{aligned}
\lambda_{\alpha'(a\cdot b)} &= \lambda_{\overline{(ab)}} \\
&= \alpha' \circ \lambda_{ab} \circ \alpha'^{-1} \\
&= \alpha' \circ \lambda_a \circ \lambda_b \circ \alpha'^{-1} \\
&= \alpha' \circ \lambda_a \circ \alpha'^{-1} \circ \alpha \circ \lambda_b \circ \alpha'^{-1} \\
&= \lambda_{\overline{a}} \circ \lambda_{\overline{b}} \\
&= \lambda_{\overline{a}\cdot\overline{b}} && = \lambda_{\alpha'(a)\cdot\alpha'(b)}
\end{aligned}$$

[III] Again, the copy of $\mathrm{Aut}(G)$ is literally $\mathrm{Aut}(G)$ while the copy of G is via λ.

and applying both sides to e_G proves that $\alpha'(a \cdot b) = \alpha'(a) \cdot \alpha'(b)$. This means α' is o.p. which was the last thing we needed to prove in order to be able to conclude that α' is an automorphism of G, as desired.

Therefore (10) holds using $g = \alpha(e_G)^{-1}$ and $\varphi = \alpha'$. ■

13 Sylow's Theorems

Sylow's Theorem's give us {put something here}

13.1 Definitions and Statement

Throughout this chapter p is a positive prime integer.
Recall Definition 10.122 back on page 581.

Definition 13.1. A *p-group* is a group whose order is p^n for some positive integer n.

Example 13.2.

- $\mathcal{D}_4$, $\mathcal{D}_8$, $B\mathcal{D}_4$, $\langle (1\ 2\ 3\ 4), (1\ 3)(5\ 6\ 7\ 8)\rangle$, $\mathbb{Z}_2$, $\mathbb{Z}_4$ are all 2–groups.
- $\mathrm{UT}_3(\mathbb{Z}_3)$, Af_3, and $\mathbb{Z}_3 \times \mathbb{Z}_3$ are 3–groups.
- $\mathrm{UT}_n\left(\mathbb{Z}_p\right)$[1], $\mathbb{Z}_{p^2}$, $\mathbb{Z}_p \times \mathbb{Z}_p$, are all p-groups.
- $\mathcal{S}_3$, $\mathcal{A}_4$, $\mathbb{Z}_{15}$ are not p-groups (for any prime p).
- If G is a p-group, then every (proper) subgroup and (non-trivial) quotient (hence the image of every homomorphism with domain or codomain G) is also a p-group. On the other hand suppose G is a group. Moreover assume that $N \trianglelefteq G$ and G/N are p–groups. Then G itself is a p-group.

Exercise 13.1. Prove that a finite, non-trivial group G is p–group iff p divides the order of every element of G.

SOLUTION

[1] We defined the more general versions of these groups in Example/Theorem 12.5 where we showed that they have order $p^{\binom{n}{2}}$.

DEFINITION 13.3. Let G be a group.

(1) A *Sylow p-(sub)group* of G is a subgroup H of G so that

- $|H|$ is a power of p,
- $|H|$ is the largest power p which divides $|G|$.

(2) Generically, a subgroup of G is called a *Sylow* subgroup (of G) if it is a Sylow p–subgroup of G for some p.

In other words, the order of a Sylow p-group has the property that p does not divide the index of H in G. In terms of the prime factorization

$$|G| = p_1^{n_1} \cdots p_m^{n_m}$$

into powers of distinct primes, a Sylow p_i-subgroup would have order $p_i^{n_i}$. Or really beating it to death, yet another way of looking at it is if we factor $|G| = p^r m$ with $r > 0$ and $p \nmid m$, than a Sylow p–subgroup of G is *any* subgroup of G of order p^r. Note that if p^r is the order of a Sylow p–subgroup of G, then p^r divides the order of G, but p^{r+1} does not. Therefore, at least in theory, G can have a subgroup of order p^r, i.e. a Sylow p–subgroup, but G cannot contain a p–subgroup which properly contains a Sylow p–subgroup.

For an integer n and prime p, recall the notation $\nu_p(n)$ from Notation 3.161 back on page 160. Note that $p^{\nu_p(n)} \mid n$ but no higher power of p divides n.

NOTATION 13.4. For a finite group G, $\nu_p(G) \stackrel{\text{def}}{=} \nu_p(|G|)$.

In other words, if $p \Big| |G|$, $\nu_p(G)$ is defined in such a way that $p^{\nu_p(G)}$ is the order of a possible Sylow p–subgroup of G.

EXAMPLE 13.5.

(1) We consider the alternating group on 5 letters, $\mathcal{A}_5$ which has order $60 = 5 \cdot 3 \cdot 2^2$. Then a Sylow 5-subgroup would have order 5 (since $\nu_5(\mathcal{A}_5) = 1$), a Sylow 3-subgroup would have order 3 (since $\nu_3(\mathcal{A}_5) = 1$), and a Sylow 2-subgroup would have order 4 (since $\nu_2(\mathcal{A}_5) = 2$). In fact the groups

$$\langle (1\,2\,3\,4\,5) \rangle \qquad \langle (1\,2\,3) \rangle \qquad \langle (1\,2)(3\,4), (1\,3)(2\,4) \rangle$$

have orders 5, 3, and 4 respectively, so they are, respectively, Sylow 5, 3, and 2–subgroups of $\mathcal{A}_5$.

(2) We've seen that

$$|\mathbf{SL}_3(\mathbb{Z}_3)| = 5616$$

(see Table 10.1) which factors as $2^4 \cdot 3^3 \cdot 13$. Thus a Sylow 2–subgroup of $\mathbf{SL}_3(\mathbb{Z}_3)$ has order 16, a Sylow 3–subgroup would have order 27, and a Sylow 13–subgroup has order 13. In fact, we've seen already that $\mathrm{UT}_3(\mathbb{Z}_3) \leqslant \mathbf{SL}_3(\mathbb{Z}_3)$ has order 27, so it is one example of a Sylow 3–subgroup of $\mathbf{SL}_3(\mathbb{Z}_3)$.

(3) More generally for p prime, we've seen that

$$\left|\mathbf{SL}_3\left(\mathbb{Z}_p\right)\right| = p^3 \cdot (p^3 - 1) \cdot (p^2 - 1).$$

Clearly p cannot divide $p^3 - 1$ or $p^2 - 1$, so a Sylow p–subgroup of $\left|\mathbf{SL}_3\left(\mathbb{Z}_p\right)\right|$ has order p^3. Since the subgroup $\mathrm{UT}_3\left(\mathbb{Z}_p\right)$ of $\mathbf{SL}_3(\mathbb{Z}_3)$ has order p^3 it indeed is a Sylow p–subgroup of $\mathbf{SL}_3\left(\mathbb{Z}_p\right)$.

(4) Even more generally then the previous example, we've seen in Exercise 10.55 that

$$\left|\mathbf{SL}_n\left(\mathbb{Z}_p\right)\right| = p^{\binom{n}{2}}\prod_{i=2}^{n}\left(p^i-1\right).$$

Clearly p cannot divide any term except for the first, $p^{\binom{n}{2}}$. Therefore the order of a Sylow p–subgroup of $\mathbf{SL}_n\left(\mathbb{Z}_p\right)$ has order $p^{\binom{n}{2}}$. However, we've seen in Example/Theorem 12.5 that this is the order of $\mathrm{UT}_n\left(\mathbb{Z}_p\right)$. Therefore $\mathrm{UT}_n\left(\mathbb{Z}_p\right)$ is a Sylow p–subgroup of $\mathbf{SL}_n\left(\mathbb{Z}_p\right)$. Incidentally, everything I've said here also works for the larger group $\mathbf{GL}_n\left(\mathbb{Z}_p\right)$ since that group has order $p-1$ times the order of $\mathbf{SL}_n\left(\mathbb{Z}_p\right)$ which contributes no additional factors of p.

(5) Let G be a group of order

$$88,179,840 = \cdot 7\cdot 5\cdot 3^9\cdot 2^7$$

Then a Sylow 2–subgroup of G would have order 2^7, a Sylow 3–subgroup would have order 3^9, a Sylow 5–subgroup would have order 5, and a Sylow 7–subgroup would have order 7. G would not have a Sylow p–subgroup for any other prime number p.

(6) Since $B\mathcal{D}_4 \leqslant B\left(\blacklozenge^3\right)$ and since $\left|B\left(\blacklozenge^3\right)\right| = 48 = 16\cdot 3$, we have $B\mathcal{D}_4$ is a Sylow 2–subgroup of $B\left(\blacklozenge^3\right)$.

(7) The groups $\langle(1\,2\,3\,4)(1\,3)\rangle$, $\langle(1\,3\,2\,4),(1\,2)\rangle$, and $\langle(1\,3\,4\,2),(2\,3)\rangle$ are each Sylow 2–subgroups of $\mathcal{S}_4$ (since they are subgroups of $\mathcal{S}_4$ of order 8) which are isomorphic to $\mathcal{D}_4$.

(8) If G is already a p-group, then the only Sylow subgroup of G is G itself.

NOTATION 13.6. Let $\mathbf{Syl}_p(G)$ be the set of Sylow p–subgroups of G and

$$n_p(G) \stackrel{\text{def}}{=} \#\left(\mathbf{Syl}_p(G)\right).$$

EXAMPLE 13.7.

(1) If p does not divide G, then Lagrange's theorem tells us that $n_p(G) = 0$.

(2) If G is a p-group, then $n_p(G) = 1$ and $n_q(G) = 0$ for all other prime numbers q.

(3) It turns out that $n_2(\mathcal{S}_4) = 3$ with the 3 Sylow 2-groups as listed in the above example. Proving this, even in the case of this simple example, is rather difficult to do "by hand" but is trivial to do using the Sylow theorems below.

Recall that a subgroup, H, of a group G is characteristic if for all automorphisms φ of G, $\varphi(H) = H$. This was denoted H **char** G. Remember that since a particular kind of automorphism is conjugation by a given element, we have that characteristic subgroups are normal.

LEMMA 13.8. *Suppose that $|H|$ is the only subgroup of a group G of order n. Then H* **char** *G. (And hence $H \trianglelefteq G$).*

Proof. For any $\varphi \in \mathrm{Aut}(G)$ we have that $\varphi(H)$ is also a group of order n. But since there is only one such group by hypothesis, we must have that

$$\varphi(H) = H.$$

Thus H must be characteristic in G. ∎

Theorem 13.9 (Sylow's Theorems). *Suppose that $p \mid |G|$. Then*

Sylow1 **(Existence)**
There exists a Sylow p–subgroup of G and every p–subgroup of G is contained inside of some Sylow p–subgroup of G.

Sylow2 **(Uniqueness)**
All Sylow p–subgroups of G are conjugate, and hence isomorphic.[II]

Sylow3 **(Number)**
And

- **(a)** $|G| \equiv_{n_p(G)} 0$; (*That is $n_p(G)$ divides $|G|$*)
- **(b)** $n_p(G) \equiv_p 1$. (*In particular, p never divides $n_p(G)$*)

Remark 13.10. Combining **Sylow1** and **Sylow2**, we see for any p–subgroup of G, $\widehat{P}$ and any Sylow p–subgroup of G, there exists a $g \in G$ so that $g\widehat{P}g^{-1} \leqslant P$. Indeed, **Sylow1** says that $\widehat{P}$ is contained inside of some Sylow p–group, $\widetilde{P}$, and **Sylow2** says that $\widetilde{P}$ is conjugate to P; meaning there exists a $g \in G$ so that $gP'g^{-1} = P$. Therefore,

$$g\widehat{P}g^{-1} \leqslant g\widetilde{P}g^{-1} = \widehat{P}$$

as promised. In some textbooks, this is stated as Sylow 2.

Remark 13.11. **Sylow3** says that $n_p(G)$ both divides $|G|$ and is relatively prime with p. This means if we factor $|G|$ as $p^{\nu_p(G)}m$, then $n_p(G)$ must divide m. This will prove to have little theoretical importance, but will serve as a great time saver!

Remark 13.12. We immediately see that **Sylow2** implies that **Sylow3.a**. This is because **Sylow2** says that the action of G, by conjugation, on its set of Sylow p–groups is transitive. But whenever any group acts transitively on any set, the size of the set divides $|G|$. This is a consequence of the orbit stabilizer theorem. Actually, since as we've seen, the stabilizer of a particular subgroup under this action is the normalizer of the given subgroup. Thus as a direct consequence of **Sylow2** and **OST**, we have

$$n_p(G) \cdot |\mathrm{N}_G(P)| = |G|$$

[II] In more detail, for any two Sylow p–subgroups of G, P and P', there exists a $g \in G$ with $gPg^{-1} = P'$.

where P is any Sylow p–subgroup of G. In particular, since P is normal in G iff its normalizer, $\mathrm{N}_G(P)$ is all of G, i.e. iff $|\mathrm{N}_G(P)| = |G|$, this proves that

A Sylow p–subgroup of G is normal iff $n_p(G) = 1$.

Indeed, one direction, the direction which says if $n_p(G) = 1$, then a Sylow p–subgroup of G is normal, is obvious without the Sylow theorems. Saying that $n_p(G) = 1$ means that there is 1 and only 1 subgroup of G of order $p^{\nu_p(G)}$ (where $p^{\nu_p(G)}$ is the order of a Sylow p–subgroup of G). Therefore, this unique subgroup of that order must be normal, even characteristic in G (c.f. Lemma 13.8).

The other direction does need **Sylow2**. For instance, in $\mathcal{D}_4$ there are a total of 5 subgroups of order 2. However, 1 of them (and only 1 of them), namely the one generated by the 180° rotation is normal in $\mathcal{D}_4$. Meanwhile there are a total of 3 subgroups of order 4 in $\mathcal{D}_4$. All three, having index 2, are normal in $\mathcal{D}_4$. However only 1 of them, namely the one generated by the 90° rotation is actually characteristic in $\mathcal{D}_4$.[III]

13.2 Some Applications Before the Proof

In these examples G is a group and P_p is a (any) Sylow p–subgroup of G. We assume the Sylow's theorems are true here to illustrate some uses.

Theorem 13.13. *Suppose that* $|G| = p^r \cdot m$ *where* $\gcd(p, m) = 1$. *Suppose further that with the exception of the number 1, no divisor of* m *is equivalent to 1 mod* p. *Then*

(1) *There exists a unique subgroup* $P \leqslant G$ *of order* p^r

(2) *The subgroup* P *is normal in* G (*It's even a characteristic subgroup* G).

(3) *Every* p*–subgroup of* G *is contained inside of* P. *In particular, every element of* G *whose order is a power of* p *is contained in* G.

Proof.

(1) By **Sylow1** there is a subgroup, P, of G whose order is p^r. Every divisor of $|G|$ is equal to $n = p^s \cdot D$ where $s \leq r$ and D is a divisor of m. If $n \equiv_p 1$, then we must have that $s = 0$ for otherwise n would be divisible by p. Moreover our assumption is that $D - 1$ is not divisible by p. Therefore the only divisor of n which is equivalent to 1 mod p is the number 1 itself. Therefore by **Sylow3** there is a unique subgroup of G of order p^r.

(2) Lemma 13.8 applies proving that P **char** G and $P \trianglelefteq G$.

(3) **Sylow1** says that every p–subgroup of G is contained inside of some Sylow p–subgroup of G. But we have just proven that P is the only such group. Thus, every p–subgroup of G must be contained inside of P.

To prove the remark about orders, not that for every element of group of order a power of a p generates a group, which happens to contain it, of order a power of p. Thus this subgroup, and hence the element itself, must be contained inside of P. ∎

By **Sylow2**, it makes sense to talk about *the* Sylow p–subgroup of G (up to isomorphism). For instance, any group of order

$$591,917,144 = 2^3 \cdot 7 \cdot 13 \cdot 23^2 \cdot 29 \cdot 53$$

must (by **Sylow1**) contain a subgroup isomorphic to *exactly* (by **Sylow2**) one of the following groups:

$$Q_8,\ \mathcal{D}_4,\ \mathbb{Z}_8,\ \mathbb{Z}_4 \times \mathbb{Z}_2,\ \text{ or } \mathbb{Z}_2^3.$$

This is because, up to isomorphism, these are the only group of order 8. It also says that if say, one group of this massive order happens to contain, say, $\mathbb{Z}_8$ and another happens to contain, say, Q_8, then those two groups cannot be isomorphic. In fact, this second group would not even be allowed to contain an element of order 8 at all! This is because such an element would generate a copy of $\mathbb{Z}_8$.

Exercise 13.2. Show that $\mathbf{SL}_3(\mathbb{Z}_2)$ does not contain an element of order 8.
Hint! Show that $\mathrm{UT}_3(\mathbb{Z}_2)$ does not contain an element of order 8.

Solution

Exercise 13.3. For p an odd prime (i.e. a prime not equal to 2), show that $\mathbf{SL}_3\left(\mathbb{Z}_p\right)$ does not contain an element of order p^2.
Hint! Recall that $\mathrm{UT}_3\left(\mathbb{Z}_p\right)$ is a Sylow p–subgroup of $\mathbf{SL}_3\left(\mathbb{Z}_p\right)$. Then use Exercise 5.41.

Solution

The hardest part of this theorem to get used to is the third part. The two parts of this often work in tandem. For instance:

Example 13.14. As a first example we use these theorems to show that there is only one group of order 15, namely $\mathbb{Z}_{15} = \mathbb{Z}_5 \times \mathbb{Z}_3$. Let G be a group of order 15 and let P_5 be a Sylow 5–subgroup of G and P_3 an Sylow 3–subgroup of G. Since $\nu_5(15) = \nu_3(15) = 1$, we have that $|P_5| = 5$ and $|P_3| = 3$. In particular, each have prime order and so

$$P_3 \cong \mathbb{Z}_3 \quad \text{and} \quad P_5 \cong \mathbb{Z}_5.$$

By **Sylow3**, the number of 3 groups of G, $n_3(G)$, divides 15 and is equivalent to 1 mod 15. But of the divisors of 15, which are $\{1, 3, 5, 15\}$ the only number which is equivalent to 1 mod 15 is 1 itself. Thus $n_3(G) = 3$ and so P_3 is the only subgroup of G of order 3. Similarly P_5 is the only subgroup of G of order 5. Therefore, by Lemma 13.8 we have that both P_3 and P_5 are normal in G. Since their orders are relatively prime, we know that $P_3 \cap P_5 = \{e_G\}$. Lastly, we must have that $G = \langle P_3, P_5 \rangle$ since

$$15 = \mathrm{lcm}(5, 3) \Big| |\langle P_3, P_5 \rangle|$$

by Lagrange's theorem. Therefore by the Direct product Recognition Theorem (Theorem 8.77 back on page 426) we must have that

$$G = P_5 \times P_3 \cong \mathbb{Z}_5 \times \mathbb{Z}_3$$[I]

as desired.

In the above example we essentially prove the following proposition:

Proposition 13.15. *Suppose $n_p(G) = 1$ for every prime divisor of the order of a finite group G. Then*

$$G = P_{p_1} \times \cdots \times P_{p_m}$$

where $p_1, \ldots, p_m$ are all prime divisors of $|G|$ and P_p is a Sylow p–subgroup of G.

Proof. Suppose that $|G| = p_1^{n_1} \cdots p_m^{n_m}$ is the prime factorization of the number $|G|$. The groups P_{p_i} exist by **Sylow1**. And are each normal by **Sylow3**.

Now for distinct prime divisors of $|G|$, p and q, we have that $P_p \cap P_q = \{e_G\}$ by Lagrange's theorem. Since P_p and P_q are each normal in G, they are also each normal in $\left\langle P_p, P_q \right\rangle$. Therefore, by the direct product recognition theorem,

$$\left\langle P_p, P_q \right\rangle \cong P_p \times P_q.$$

[I] So G is isomorphic to $\mathbb{Z}_{15}$.

Using this as the base case for an induction, we now prove that for every $i \leq m$,

$$\left\langle P_{p_1}, \ldots, P_{p_i} \right\rangle \cong P_{p_1} \times \cdots \times P_{p_i}.$$

The desired result is the special case of this with $i = m$. Indeed, if we assume the result holds for all integers $i - 1$, then

$$\left\langle P_{p_1}, \ldots, P_{p_{i-1}} \right\rangle \cong P_{p_1} \times \cdots \times P_{p_{i-1}} \Rightarrow \left| \left\langle P_{p_1}, \ldots, P_{p_{i-1}} \right\rangle \right| = p_1^{n_1} \cdots p_{i-1}^{n_{i-1}}$$

implies that the order of $\left\langle P_{p_1}, \ldots, P_{p_{i-1}} \right\rangle$ is relatively prime with P_{p_i}. Hence, again by Lagrange, this implies that P_{p_i} meets $\left\langle P_{p_1}, \ldots, P_{p_{i-1}} \right\rangle$ only trivially. Moreover both are clearly normal in G. Therefore, by the direct product recognition theorem, we know that

$$\left\langle P_{p_1}, \ldots, P_{p_{i-1}}, P_{p_i} \right\rangle \cong \left\langle P_{p_1}, \ldots, P_{p_{i-1}} \right\rangle \times P_{p_i} \cong \left(P_{p_1} \times \cdots \times P_{p_{i-1}} \right) \times P_{p_i}$$

as desired. ∎

PROPOSITION 13.16. *Suppose that $p \left| |G|$, but p^2 does not. Then there are $n_p \cdot (p-1)$ elements of order p in G (where $n_p = n_p(G)$, the number of Sylow p–subgroups of G).*

PROOF. A Sylow p–subgroup of G must have order p by assumption, and therefore must have $p - 1$ elements of order p and the identity in it. Conversely, every element of order p in G generates a subgroup of order p, and hence a Sylow p–subgroup. Now two distinct Sylow p–subgroups must have trivial intersection since there intersection must have order a divisor of p which doesn't equal p itself since they are distinct. Therefore the $n_p(G)$ Sylow p–subgroups partition the set of all elements of order p in G into $p - 1$ element groupings. Hence the result. ∎

More generally, if we know that $P \in \mathbf{Syl}_p G$ is cyclic or order p^r, then we know that there are exactly

$$n_p \cdot \vartheta(p^r) = n_p \cdot p^{r-1} \cdot (p-1)$$

elements of order p^r in G.

EXAMPLE 13.17. I claim that $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ contains $(p^2 - 1)$ elements of order p. We have shown that $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ has order $(p-1)\,p\,(p+1)$. Thus the order of a Sylow p–subgroup of $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ is p. Now an example of a Sylow p–subgroup of $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ is $\langle \Omega \rangle$, where as usual Ω is the matrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

taken mod p, and so

$$\langle \Omega \rangle = \left\{ \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} : a \in \mathbb{Z}_p \right\}.$$ [II]

This subgroup is not obviously not normal in $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$. For instance,

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \notin \langle \Omega \rangle$$

[II] Of course this is $\mathrm{UT}_3\left(\mathbb{Z}_p\right)$, and since we've shown that this is always a (the) Sylow p–subgroup of $\mathbf{SL}_n\left(\mathbb{Z}_p\right)$, this shouldn't be a suprise!

proves this. In fact, if we use Example/Theorem 8.55, we see that $\mathrm{UT}_n\left(\mathbb{Z}_p\right)$ is $\mathrm{ST}_n\left(\mathbb{Z}_p\right)$, which in this case it's the set of matrices

$$\left\{\begin{pmatrix} a & b \\ 0 & \frac{1}{a} \end{pmatrix} : a, b \in \mathbb{Z}_p, a \neq 0\right\}.$$

Therefore, there must be

$$\frac{(p+1)\,p\,(p-1)}{p\,(p-1)} = p+1$$

Sylow p–subgroups of $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$. So we can apply the theorem to conclude that there are $(p+1)(p-1)$ elements of order p in $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$. For instance, there are a total of $8 \cdot 6 = 48$, 2 by 2 matrices with coefficients in $\mathbb{Z}_7$ of determinant 1 and order 7.

Example 13.18. We show that any group of order 30 has an element of order 15. By the previous example, we can do this by showing that G has a subgroup of order 15. Assume that the Sylow 5-subgroup and the Sylow -subgroups of G are not normal. By Sylow 3, there must therefore be 6 Sylow 5-groups and 10 Sylow 3-subgroups of G. By the theorem this would imply that there are a total of 20 elements of order 3 and 24 elements of order 5 in G. Therefore at least one of P_5 or P_3 must be normal. Therefore, P_3P_5 is a group which has order 15 and hence the element of order 15 in G.

Example 13.19. In $\mathcal{S}_6$, which has order

$$720 = 6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5 \cdot 3^2 \cdot 2^4,$$

we have that the order of a Sylow 3–subgroup is 9. Indeed, we can see directly that an example of a subgroup of $\mathcal{S}_6$ of order 9, hence an example of a Sylow 3–subgroup of $\mathcal{S}_6$, is

$$P_3 := \langle (1\,2\,3), (4\,5\,6) \rangle$$

(which is isomorphic to $\mathbb{Z}_3 \times \mathbb{Z}_3$). Checking divisors (again, only need to check the divisors of $5 \cdot 2^4$) we see there can be either

$$4, 10, 16, \text{ or } 40$$

such subgroups (i.e. $n_3(\mathcal{S}_6) = 4$,10,16, or 40). This is because these are the only divisors of 720 which are equivalent to 1 mod 3. We can figure out exactly which it is in several ways:

- Perhaps the fastest way is to note that the normalizer of P_3 is

  $$\langle (1\,2\,3), (1\,2), (4\,5\,6), (4\,5), (1\,4)(2\,5)(3\,6) \rangle$$

  which has order $3! \cdot 3! \cdot 2 = 72$ (The first 2 pairs of 2 generators each generate a subgroup isomorphic to $\mathcal{S}_3$ and the last generator exchanges these two copies.) Therefore, there are a total of

  $$\frac{6!}{3! \cdot 3! \cdot 2} = 5 \cdot 2 = 10$$

  Sylow 3–subgroups of $\mathcal{S}_6$.

- Using the formula from Theorem 10.113 back on page 565 we know that there are

  $$\frac{6!}{3^2 \cdot 2} = 5 \cdot 4 \cdot 2 = 40 \tag{1}$$

  $\underline{3}\,\underline{3}$–cycles in $\mathcal{S}_6$. Note that in P_3 there are a total of 4 such elements, namely

  $$(1\,2\,3)(4\,5\,6),\ (1\,3\,2)(4\,5\,6),\ (1\,2\,3)(4\,6\,5),\ (1\,3\,2)(4\,6\,5).$$

  Next note that any two distinct Sylow 3–subgroups of $\mathcal{S}_6$ must meet trivially. This is due to the following two facts:

(1) The only Sylow 3–subgroup of $\mathcal{S}_6$ which contains the $\underline{3}\,\underline{3}$–cycle $(a\ b\ c)(x\ y\ z)$ is

$$\langle (a\ b\ c), (x\ y\ z) \rangle .$$

(2) A $\underline{3}$–cycle commutes with another $\underline{3}$–cycle iff they are powers of a common $\underline{3}$–cycle or if they are disjoint.

Therefore, there are a total of $n_3(\mathcal{S}_6) \cdot 4$ $\underline{3}\,\underline{3}$–cycles in $\mathcal{S}_6$. Checking this against (1) shows (once again) that $n_3(\mathcal{S}_6) = 10$.

Example 13.20. (Groups of order 1309) The number 1309 factors as $7 \cdot 11 \cdot 17$.

- Mod 7, $11 = 4$ and $17 = 3$ and their product is equivalent to $12 \equiv_7 5$, none of which is equivalent to 1 mod 7. Therefore $n_7(G) = 1$ for any group G of order 1309.
- Mod 11, $7 = 7$ and $17 = 6$ and their product is $42 \equiv_{11} 9$. Therefore $n_{11}(G) = 1$ for any group G of order 1309.
- Mod 17, none of 7, 11, nor $7 \cdot 11 = 77 \equiv_{17} 9$ are equivalent to 1 mod 17. Therefore we also have $n_{17}(G) = 1$ for any group G.

Putting these three facts together, and the fact that any group of prime order is isomorphic to the cyclic group of that order, we have proven that any group of order 1309 is isomorphic to

$$\mathbb{Z}_7 \times \mathbb{Z}_{11} \times \mathbb{Z}_{17}.$$

Maple Example 13.21. Let's write a code to save time on some calculations. The input of the code is a natural number n. The output is a list of pairs $[p, A]$ where p runs through the prime divisors of n and A is the set of divisors of n which are equivalent to 1 mod p. In other words, A is the set of possible values of $n_p(G)$ where G is a group of order n. The code is as follows:

```
> with(numtheory);
> FF:=proc(n)
   local L, flag, i,p, flag2;
   flag:={};
   flag2:={};
   L:=[op(divisors(n))];
   for p in factorset(n) do
   for i in L do
   if(i mod p=1) then
   flag2:=flag2 union {i}; fi;
   od;
   flag:=flag union {[p,flag2]};
   flag2:={};
   od;
   return(flag);
 end;
```

(1) For instance, if I apply this to $n = 5005 = 5 \cdot 7 \cdot 11 \cdot 13$ I get the output

$$\{[5, \{1, 11, 91, 1001\}], [7, \{1, 715\}], [11, \{1\}], [13, \{1\}]\}.$$

This indicates that:

- There are

  $$1, 11, 91, \text{ or } 1001$$

  Sylow 5–subgroups of a group of order 5005.

- There are

  $$1 \text{ or } 715$$

  Sylow 7–subgroups of a group of order 5005.
- There can only be

  $$1$$

  Sylow 11–subgroups of a group of order 5005. In particular, any group of order 5005 has a normal subgroup of order 11.
- There can only be

  $$1$$

  Sylow 13–subgroups of a group of order 5005. In particular, any group of order 5005 has a normal subgroup of order 13.
- We also have that $\langle P_{13}, P_{11} \rangle$ is cyclic (why?) and has order $13 \cdot 11 = 143$. In particular, we have proven that any group of order 5005 has an element of order 143.

(2) Sometimes **Sylow3** doesn't give us too much. For instance, plugging $360 = 2^3 \cdot 3^2 \cdot 5$ into the `FF` gives

```
{[2, {1, 3, 5, 9, 15, 45}], [3, {1, 4, 10, 40}], [5, {1, 6, 36}]}.
```

This does not tell us a whole lot!

MAPLE EXERCISE 13.4.

SOLUTION

(1) Use `MAPLE` and the program above to compute the possible values of $n_7(G)$ where G is a group of order $7! = 5040$.

(2) As a particular example of a group of order 7!, compute $n_7(\mathcal{S}_7)$.

EXERCISE 13.5. Use the Sylow's theorems to prove that there is only one group of order 35, namely $\mathbb{Z}_{35}$.

SOLUTION

EXERCISE 13.6. What are the isomorphism types of a Sylow 2-subgroup, Sylow 3-subgroup, and a Sylow 7-subgroup of $\mathcal{A}_5$?
Hint! To figure out what a Sylow 2-subgroup is you might find Exercise 9.43 back on page 478 useful.

SOLUTION

EXERCISE 13.7. Show that $\mathcal{D}_4$ contains non-isomorphic (hence non-conjugate) subgroups of order 4. Show that $\mathcal{D}_4$ also contains isomorphic subgroups of orders both 4 and 2 which are not conjugate. Why don't these facts contradict **Sylow2**?

SOLUTION

13.3 Applications w/Semi–Direct Products*

This optional section if for those who've read about semi-direct products, Section 12.3.

Example 13.22. For many numbers n, up to isomorphism, every group is a semi–direct products. Using Sylow's theorems one can show this quite quickly. Therefore classifying groups of orders of such numbers therefore reduces to classifying semi–direct products, which is not also an easy task, but at least it's manageable: it amounts to characterizing conjugacy classes of automorphisms.

(1) Let's classify groups whose order is the product of just two primes. Let G be a group of order $n = p\,q$ with p and q prime numbers w/$p < q$. There are two possibilities:

Case 1 $q \neq [p]1$, e.g. $n = 15 = 5 \cdot 3, n = 33 = 11 \cdot 3, n = 35 = 7 \cdot 5, n = 161 = 7 \cdot 23$.

Case 2 $q \equiv_p 1$, e.g. $n = 21 = 7 \cdot 3, n = 39 = 13 \cdot 3, n = 55 = 11 \cdot 5, n = 155 = 31 \cdot 5$.

The first case, **Case 1**, is easier, and a simple generalization of 13.14 shows that

Theorem 13.23. *Every group fitting into* **Case 1** *is isomorphic to* $\mathbb{Z}_p \times \mathbb{Z}_q \cong \mathbb{Z}_{pq}$. ∎

Therefore, up to isomorphism, there is only one group of orders $n = 15$, $n = 33$, $n = 35$, $n = 161$, etc. In fact, one could show with only a bit more effort that for an integer m, if there exists only one group of order m (up to isomorphism), then m is a prime or a product of two primes of this form.

Case 2 is more interesting. We know that the Sylow q group, P_q, is still normal in G. This is due to the fact that since p is assumed to be strictly less than q, there's no way that it could be equivalent to 1 mod q. However, since we **are** assuming that $q \equiv_p 1$, it's entirely possible that the Sylow p–subgroup, P_p, of G is not normal in G. That's ok because we can apply the semi-direct product recognition theorem, Theorem 12.14, to still be able to conclude that

$$G \cong \mathbb{Z}_q \rtimes_\bullet \mathbb{Z}_p.$$

Clearly the trivial action is possible and in which case $G \cong \mathbb{Z}_{pq}$. For starters, by Corollary 9.49 $\mathrm{Aut}\left(\mathbb{Z}_q\right)$ is isomorphic to $\mathcal{U}(\mathbb{Z}_q)$. Moreover, since q is a prime, the group $\mathcal{U}(\mathbb{Z}_q)$ is cyclic (see Theorem 5.75) and has order $q-1$. Therefore, since by assumption $p \mid (q-1) = \left|\mathcal{U}(\mathbb{Z}_q)\right|$, Corollary **??** applies to show there is a unique subgroup of $\mathcal{U}(\mathbb{Z}_q)$ of order p. Hence there is only one possible non-trivial image of a permutation representation

$$\mathbb{Z}_p \to \mathcal{U}(\mathbb{Z}_q).$$

This is enough {should I say more?} to be able to conclude that up to isomorphism there is a unique non-trivial (i.e. one which does not lead to a direct product) semi–direct product of the form $\mathbb{Z}_q \rtimes_\bullet \mathbb{Z}_p$. Thus we have classified groups of order $p\,q$ in this second (and final) case:

Theorem 13.24. *Every group fitting into* **Case 2** *is isomorphic to either*

(a) (**Abelian Case**)

$$\mathbb{Z}_{pq}$$

(b) (**Non-abelian Case**)

$$\mathbb{Z}_q \rtimes_m \mathbb{Z}_p.$$

(*So* $yxy^{-1} = x^m$ *where* m *is an element of multiplicative order* p *in* $\mathcal{U}(\mathbb{Z}_q)$, *and* x *is a generator of* $\mathbb{Z}_q$ *and* y *is a generator of* $\mathbb{Z}_p$.)

Therefore, up to isomorphism, there are two groups of order $n = 21$, $n = 39$, $n = 55$, $n = 155$, etc. To increase your intuition, let's kick around the non-abelian group of order 21, which we call G_{21}. It has two generators x and y, with $|x| = 7$ and $|y| = 3$. Now $[2]_7$ has multiplicative order 3 ($2^3 = 8 = 1$ mod 7) we can take $yxy^{-1} = x^3$. (Up to isomorphism these three facts are enough to determine this group.) In fact, in $\mathcal{S}_7$ the normalizer of the group generated by the 7–cycle $(1\,2\,3\,4\,5\,6\,7)$ is

$$\mathrm{N}_{\mathcal{S}_7}\langle(1\,2\,3\,4\,5\,6\,7)\rangle = \langle(1\,2\,3\,4\,5\,6\,7), (1\,3\,2\,6\,4\,5)\rangle.^{1}$$

It's plain to see that this group has order 42 (second generator normalizes first, and powers of the two generators do not intersect). The second generator listed is a 6–cycle and, as such, is odd. Therefore, in $\mathcal{A}_7$ the normalizer is

$$\mathrm{N}_{\mathcal{S}_7}\langle (1\,2\,3\,4\,5\,6\,7)\rangle = \langle (1\,2\,3\,4\,5\,6\,7), (2\,3\,5)(4\,7\,6)\rangle$$

which is (isomorphic to) our non–abelian group of order 21, G_{21}. In fact, this process always works to give us a concrete way of looking at our group of order pq (with $p \equiv_q 1$) as a permutation group.

Exercise 13.8. List the conjugacy classes and write down the order table for the group G_{21} just discussed.

Solution

(2) Let's classify the groups of order 18. In such a group, which we call (what else?) G there must be a unique Sylow 3 group, P_3 (must have order 9). Also, we let P_2 be a Sylow 2-group (must have order 2 here). The usual argument shows us that $G = P_3P_2$ and $P_2 \cap P_3$. Therefore any group of order 9 must be a semi–direct product of P_2 and P_3. By the classification of groups of order p^2 we must have that $P_3 \cong \mathbb{Z}_9$ or $P_3 \cong \mathbb{Z}_3 \times \mathbb{Z}_3$. Moreover, since there is only one group of order 2 we must have that $P_2 \cong \mathbb{Z}_2$. The automorphism group of $\mathbb{Z}_9$ has order 6 and is isomorphic to $\mathcal{U}(\mathbb{Z}_9) \cong \mathbb{Z}_6$. Therefore, there is only one element of order 2, which is the element $[-1]_9$. This corresponds to the automorphism which takes $[m]_9 \to [-m]_9$. Therefore, a homomorphism

$$\mathbb{Z}_2 \to \mathrm{Aut}(\mathbb{Z}_9) \qquad (\mathrm{Aut}(\mathbb{Z}_9) \cong \mathcal{U}(\mathbb{Z}_9) \cong \mathbb{Z}_6)$$

must take $[1]_2$ to either $\mathbf{1}_{\mathbb{Z}_9}$ or to $-\mathbf{1}_{\mathbb{Z}_9}$. The first semi direct product gives us $\mathbb{Z}_{18} \cong \mathbb{Z}_9 \times \mathbb{Z}_2$ and the second gives us $\mathcal{D}_9$.

In the second case, $P_2 \cong \mathbb{Z}_3^2$ and $\mathrm{Aut}\left(\mathbb{Z}_3^2\right) \cong \mathbf{GL}_2(\mathbb{Z}_3)$. There are three conjugacy classes of elements of whose square is the identity in $\mathbf{GL}_2(\mathbb{Z}_3)$. They are represented by

$$\mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{R}_1 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \mathbf{R}_2 := \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

To confuse the hell out of you[II] write $P_2 = \langle t \rangle$, multiplicatively, and $P_3 = \langle x, y \rangle$ additively, where $|t| = 2$, $|x| = |y| = 3$ and x and y commute.

In the case that t acts via $\mathbf{I}_2$, the trivial action

$$t \bullet (x, y) = (x, y),$$

of course we'd have that $P_2 \Rightarrow P_3 \cong \mathbb{Z}_3^2 \times \mathbb{Z}_2 = \mathbb{Z}_3 \times \mathbb{Z}_6$. If t acts on P_3 via $\mathbf{R}_1$, so

$$t \bullet (x, y) = (-x, y),$$

we have that $P_2 \Rightarrow P_3 \cong \mathbb{Z}_3 \times \mathcal{S}_3$. In the third case, where t acts via $\mathbf{R}_2$, so

$$t \bullet (x, y) = (-x, -y),$$

we end up with a group which is isomorphic to the fiber product

$$\mathcal{S}_3 \times_{(\varepsilon,\varepsilon)} \mathcal{S}_3 = \{(\sigma, \tau) \in \mathcal{S}_3 \times \mathcal{S}_3 : \varepsilon(\sigma) = \varepsilon(\tau)\}.$$

Since the three actions are not conjugate, we can apply a theorem from the section on semidirect products to conclude these three groups are not isomorphic. There, in conclusion we have:

[I] The 6–cycle corresponds to multiplication by 2 in $\mathbb{Z}_7$ (the numbers are shifted up by one since $[0]_7$ is the "first" element of $\mathbb{Z}_7$). In fact it's not hard to see that the normalizer of an p–cycle in $\mathcal{S}_p$ is generated by the original p–cycle and a $p-1$–cycle which represents multiplication by a generator of the cyclic group $\mathcal{U}(\mathbb{Z}_p)$ of order $p-1$. This is the $\mathrm{Hol}(\mathbb{Z}_7)$ group constructed in the last chapter.

[II] Of course there's a good reason why, but you'll see it after a minute or so. Sorry!

Every group of order 18 is isomorphic to $\mathbb{Z}_{18}$, $\mathbb{Z}_3 \times \mathbb{Z}_6$, $\mathcal{D}_9$, $\mathbb{Z}_3 \times \mathcal{S}_3$, or $\mathcal{S}_3 \times_{(\varepsilon,\varepsilon)} \mathcal{S}_3$.

Writing these as semi direct products they are

$$\mathbb{Z}_9 \times \mathbb{Z}_2 \qquad \mathbb{Z}_3^2 \times \mathbb{Z}_2 \qquad \mathbb{Z}_9 \rtimes_{a \mapsto a^{-1}} \mathbb{Z}_2$$
$$\mathbb{Z}_3^2 \rtimes_{\mathbf{R}_1} \mathbb{Z}_2 \qquad \mathbb{Z}_3^2 \rtimes_{\mathbf{R}_2} \mathbb{Z}_2.$$

More generally, this holds for groups of order $2p^2$ where p is an odd prime[III]. Namely, there are 5 groups that order which correspond to one of the above semi–direct products adapted to the more general case.

(3) (Harder!) In this example we consider groups of order $100 = 5^2 \cdot 2^2$. Let G denote an arbitrary group of order 100. Since neither 2 nor 4 are equivalent to 1 mod 5 we know that the Sylow 5–subgroup of any group of order 100 must be normal. Moreover, since both 5^2 and 2^2 are the square of a prime there are two possibilities each for the isomorphism type of the Sylow p–subgroups of G:

- $P_2 \cong \mathbb{Z}_4$ or $\cong \mathbb{Z}_2^2$
- $P_3 \cong \mathbb{Z}_9$ or $\cong \mathbb{Z}_3^2$

EXERCISE 13.9. Let G be the group

$$\mathbb{Z}_3^2 \rtimes_{\mathbf{R}_2} \mathbb{Z}_2 \qquad \text{(which is} \cong \mathcal{S}_3 \times_{(\varepsilon,\varepsilon)} \mathcal{S}_3\text{)}$$

just defined above.

SOLUTION

(1) Turn[IV] G into a subgroup of $\mathcal{S}_9$ by having it act on 9 things. (As usual, simply write down which permutations the three generators correspond to unless you really feel like writing down where all 18 elements end up.)

Please be clear what dictionary you are using since everybody will probably end up with different answers to this question, depending on how they number the 9 things you have G act on.

(2) Draw out the conjugation diagram based on the action of the three generators (called x, y, and t in the example).

(3) Write down the order table for G (that is make a listing of the number of elements of a given order which appear in G.)

EXERCISE 13.10. Prove that any group of order $1463 = 19 \cdot 11 \cdot 7$ is cyclic.
Hint! Start w/P_7, then build up to a group N of order 77 which is normal by pulling back an index 19 subgroup of the quotient G/P_7. Prove that N is cyclic and use it to classify semi–direct products of the form $N \rtimes \mathbb{Z}_{19}$ by showing that $\mathbb{Z}_6 \times \mathbb{Z}_{10}$ contains no elements of order 19.

SOLUTION

[III] Even in the case that p is not odd, i.e. in the case that $p = 2$, so $2p^2 = 8$, coincidently there are still also 5 groups. However, in this case one of the groups, Q_8, is not a semi-direct product of any pair of smaller groups. Moreover in this case there are 3 abelian and 2 non-abelian, but in the case that p were prime there are 2 abelian and 3 non-abelian.

[IV] Yeah, yeah ... by this I really mean find a monomorphism $G \to \mathcal{S}_9$, or equally rigorously, I should have said, find a subgroup of $\mathcal{S}_9$ which is isomorphic to G.

The following theorem gives us yet another way of proving that two groups are not isomorphic.

THEOREM 13.25. *Any two Sylow p–subgroups of isomorphic groups are isomorphic.*

PROOF. Suppose G and $\dot{G}$ are isomorphic finite groups.

(1) Any two Sylow p–subgroups of a single group G are conjugate, and hence isomorphic by **Sylow2**.

(2) Since whether or not a particular subgroup of a given (finite) group is a Sylow p–group is determined solely on the order of the subgroup, an isomorphism of G to $\dot{G}$ must take a Sylow p–subgroup of G to a Sylow p–subgroup of $\dot{G}$. Moreover it must do it isomorphically since it is the restriction of an isomorphism.

Combining these two results, and using transitivity of $\cong$, we see the result: Any Sylow p–subgroup of G must be isomorphic to any Sylow p–subgroup of $\dot{G}$. ■

EXAMPLE 13.26. Checking out groups of order 24, both $B\mathcal{D}_6$ and $B\left(\blacktriangleright^3\right)$ contain a Sylow 2–subgroup isomorphic to $\mathcal{Q}_8$ and a Sylow 3–subgroup isomorphic to $\mathbb{Z}_3$ (in the first case χ_6^4 generates the Sylow 3-subgroup and $\mathbf{i} = \chi_6^3$ and $\mathbf{j}$ generate $\mathcal{Q}_8$ and we have seen this many times in the second case). However in the first case the $\mathcal{Q}_8$ isn't normal, but the $\mathbb{Z}_3$ is while in the second case the opposite is true: the $\mathcal{Q}_8$ is normal, but the $\mathbb{Z}_3$ is not. The group $\mathcal{S}_4$ has $\mathcal{D}_4$ and $\mathbb{Z}_3$ as its Sylow subgroups, neither of which are normal. The group $\mathcal{D}_{12}$ has the same (isomorphism types) of Sylow subgroups, in this case the $\mathbb{Z}_3$ is normal while the $\mathcal{D}_4$ is not. Let's make a table which contains this information and more:

Group	$\mathcal{S}_4$	$B\left(\blacktriangleright^3\right)$	$B\mathcal{D}_6$	$\mathcal{D}_{12}$	$\mathcal{Q}_8 \times \mathbb{Z}_3$	$\mathcal{D}_4 \times \mathbb{Z}_3$	$\mathcal{S}_3 \times \mathbb{Z}_4$	$\mathcal{S}_3 \times \mathbb{Z}_2^2$
P_2	$\mathcal{D}_4$	$\mathcal{Q}_8$	$\mathcal{Q}_8$	$\mathcal{D}_4$	$\mathcal{Q}_8$	$\mathcal{D}_4$	$\mathbb{Z}_2 \times \mathbb{Z}_4$	$\mathbb{Z}_2^3$
n_2	3	1	3	3	1	1	3	3
P_3	$\mathbb{Z}_3$	$\mathbb{Z}_3$	$\mathbb{Z}_3$	$\mathbb{Z}_3$	$\mathbb{Z}_3$	$\mathbb{Z}_3$	$\mathbb{Z}_3$	$\mathbb{Z}_3$
n_3	4	4	1	1	1	1	1	1

Since no two groups listed have the same column, no two groups listed are isomorphic.

EXAMPLE 13.27. I claim that $\mathcal{Q}_8 \not\leq \mathcal{S}_5$ (I really mean that it's not isomorphic to a subgroup of $\mathcal{S}_5$ since it's obviously not literally a subset of $\mathcal{S}_5$). If it were, then since Sylow 2–subgroup of $\mathcal{S}_5$ must have order 8, $\mathcal{Q}_8$ would be isomorphic to this Sylow 2-subgroup of $\mathcal{S}_5$. But $\mathcal{D}_4$ is isomorphic to a subgroup of $\mathcal{S}_4$ and hence $\mathcal{S}_5$. Therefore a Sylow 2-subgroup of $\mathcal{S}_5$ must be isomorphic to $\mathcal{D}_4$ which is not isomorphic to $\mathcal{Q}_8$.

EXERCISE 13.11. Show that a Sylow 2 subgroup of $\mathcal{S}_7$ is isomorphic to $\mathcal{D}_4 \times \mathbb{Z}_2$. Show that $\mathcal{Q}_8$ is not a subgroup of this group and conclude that $\mathcal{Q}_8$ is also not a subgroup of $\mathcal{S}_7$. (It is however a subgroup of $\mathcal{S}_8$ by Cayley's theorem).

SOLUTION

13.3.1 CLASSIFICATION OF GROUPS OF ORDER 12

Let G be a group of order 12. A Sylow 2-subgroup, P_2 of G must have order 4, and a Sylow 3-subgroup, P_3, of G must have order 3. If P_3 is normal, it must be malnormal, and hence we would have a monomorphism[V] $G \to \mathcal{S}_{\#G/P_3} = \mathcal{S}_4$. Therefore the image is isomorphic to G and has order 12. But it is not hard to see that the only subgroup of order 12 in $\mathcal{S}_4$ is $\mathcal{A}_4$. Therefore we have that if $G \not\cong \mathcal{A}_4$, then $P_3 \trianglelefteq G$. So we now assume that P_3 is normal in G. Since P_3 has order 3 it is generated by some element t of

[V] Recall, a monomorphism is simply a lazy mans ways of saying an injective (1 to 1) homomorphism.

order 3. Therefore, the action of $P_2 \circlearrowright G$ by conjugation restricts to an action of P_2 on $P_3 \setminus \{e_G\}$. This action cannot possibly be faithful since P_2 has 4 elements, but $\mathcal{S}_{P_3\setminus\{e_G\}}$ has only 2. We therefore have several possibilities:

	1	2	3	4
P_2	$\mathbb{Z}_4$	$\mathbb{Z}_4$	$\mathbb{Z}_2^2$	$\mathbb{Z}_2^2$
$\mathrm{Ker}(\bullet)$	$\mathbb{Z}_2$	$\mathbb{Z}_4$	$\mathbb{Z}_2$	$\mathbb{Z}_2^2$

These possibilities arise from the fact that there are only two groups of order 4. **Case 2** and **Case 4** have $P_2 = \mathrm{Ker}(\bullet)$ which in turn implies that the group is abelian. This is because saying that $\mathrm{Ker}(\bullet) = P_3$ is "really" saying that every element of P_2 commutes with every element of P_3, and since they (P_2 and P_3 that is) together generate G, G must be abelian. In **Case 2** the group G is $\mathbb{Z}_4 \times \mathbb{Z}_3 \cong \mathbb{Z}_{12}$ and in **Case 4** G is $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \cong \mathbb{Z}_2 \times \mathbb{Z}_6$.

Now we tackle the first and third cases. For starters, it is worth noting that since in these cases, the kernel of $\bullet$ isn't all of P_2 at least one element of P_2 does not commute with every element of P_3. As such the resulting group G is not abelian in these cases. Let's be a bit more specific here. In either of the remaining cases, we denote the generator of P_3 by t. Such a generator exists since $P_3 \cong \mathbb{Z}_3$ since this is the only group of order 3.

In **Case 2**, we denote the generator (or order 4) in P_2 by a. In this scenerio we must have that a^2 commutes with t and is therefore central, but a does not commute with t. Therefore we must have that $ata^{-1} = t^{-1}$. Together the element a^2b has order 6 which we denote by b. In summary:

- $G = \langle a, b \rangle$
- $a^2 = b^3$
- $aba^{-1} = b^{-1}$
- $|b| = 6$

It is easy to see that these facts determine the multiplication table of G completely and this group is isomorphic to $B\mathcal{D}_2$ via

$$a \mapsto \mathbf{j} \quad \text{and} \quad b \mapsto \chi_6 = e^{\frac{\pi}{3}\mathbf{i}} = \cos\frac{\pi}{3} + \sin\frac{\pi}{3}\,\mathbf{i} = \frac{1}{2} + \frac{\sqrt{3}}{2}\mathbf{i}.$$

Exercise 13.12. Write down the entire table for G using only the 4 facts listed above. Conclude this group is isomorphic to $B\mathcal{D}_2$ by comparing tables.

SOLUTION

A similar logic shows that in **Case 3** we have that $G \cong \mathcal{D}_6$. Indeed, we can take generators of P_2 to be two elements, a and y, each of order 2 only one of which, say y, commuting with the generator t of P_3. Then together $b := yt$ has order 6. Then we have

- $G = \langle a, b \rangle$
- $|a| = 2$
- $aba^{-1} = b^{-1}$
- $|b| = 6$

which by Theorem 10.81 is isomorphic to $\mathcal{D}_6$.

EXERCISE 13.13. Show that no two of these groups are isomorphic.

SOLUTION

In summary we have
There are 5 groups of order 12, 2 abelian, $\mathbb{Z}_2$ and $\mathbb{Z}_6 \times \mathbb{Z}_2$, and 3 non-abelian $\mathcal{A}_4$, $\mathcal{D}_6$, and $B\mathcal{D}_2$.

13.4 PROOFS

Now we get into the actual proof of the Sylow theorems. They will extensively use the **OPST** theorem which we encountered in Lemma 10.125 back on page 582

13.4.1 LEMMAS

There are some slightly technical lemmas here all of which have the same idea, namely that for a subgroup a maximal p–subgroup, that the only things that have anything to do with p-group which normalizes the given p-group must be in the p-group itself.

DEFINITION 13.28. A p–subgroup P of G is *not maximal* if there exists a p–subgroup, $P' \leqslant G$, with P properly contained in P'. It is maximal otherwise. We denote the set of all maximal p–subgroups of G by $\mathcal{S}\mathrm{yl}_p(G)$.

The set of p–subgroups is a non-empty (by Cauchy's lemma there exists an element of order p and the group generated by that element does the trick), finite (since it is contained in the power set of a finite set, namely the group G itself) poset, and maximal p–subgroups of G are simply maximal elements in this poset. Therefore such maximal p–groups exist.

If in fact a Sylow p–subgroup of G exists, which we are in the process of proving that they do, then it must be maximal p–subgroup of G (by Lagrange's theorem). Since the converse to Lagrange's theorem is false, in general, there is no reason to expect that all maximal p-groups are Sylow p–groups. However, it will turn out that this is indeed the case and this is our first goal: to prove that

$$\mathbf{Syl}_p(G) = \mathcal{S}\mathrm{yl}_p(G)$$

It's worth pointing out that if $g \in G$ and P is a maximal p–subgroup of G that gPg^{-1} is as well. For starters this conjugate is also a p-group since conjugate groups are isomorphic and hence have the same order. Now if $gPg^{-1} \leqslant P'$ where P' is a p-group, then we would have that $P \leqslant g^{-1}P'g$ and the maximality of P guarantees that $P = g^{-1}P'g$ and so $gPg^{-1} = P'$ and gPg^{-1} is maximal.

LEMMA 13.29. *Suppose that $P \in \mathcal{S}_p(G)$. Then*

(1) *If g is in a p–subgroup of G and $g \in \mathrm{N}_G(P)$ (i.e. if $gPg^{-1} = P$), then $g \in P$.*[1]

(2) *The prime number p does not divide the index of P in its normalizer. That is*

$$p \nmid \#\left(\mathrm{N}_G(P)/P\right).$$

PROOF.

[1] In Symbols, this means that for maximal p–subgroups of G, P and P', we have that $P \cap \mathrm{N}_G(P') = P \cap P'$.

(1) Let g and P be as in the statement. Assume that g normalizes P. Let's denote the cyclic group $\langle g \rangle$ as H. Then HP is a group whose order is

$$\frac{|H| \cdot |P|}{|H \cap P|}$$

by Proposition 10.102. Since $g \notin P$ we must have that $g \neq e$ and hence, by Lagrange's theorem, must have order a positive power of a the prime p. Moreover, since $g \notin P$ we have that the order of the intersection, $|H \cap P|$, must be a smaller power of p than the order of P. Hence we must have that HP is a p-group properly containing P. This contradicts the assumption that P is maximal.

(2) Suppose that p divides this index. Since $P \trianglelefteq \mathrm{N}_G(P)$ we have that $\mathrm{N}_G(P)/P$ is a group whose order is divisible by p. By Cauchy's lemma, this group must have an element of order p. This element generates a cyclic group of order p in this quotient and whose inverse image has order $p \cdot |P|$, which is a power of p, and properly contains the group P. This contradicts the maximality assumption on P.

■

COROLLARY 13.30. *For any maximal p–subgroup of G, the restricted action of any $P \circlearrowleft \mathcal{Syl}_p(G)$ has*

$$\left(\mathcal{Syl}_p(G)\right)^P = \{P\}.$$

PROOF. By closure, we have that $gPg^{-1} = P$ for any $g \in P$. Therefore we have that $P \in \left(\mathcal{Syl}_p(G)\right)^P$. Now we show that this is the only element there. Let P and P' be distinct maximal p–subgroups of G. Then, since they are each maximal and dinstint, neither can be contained in the other, and we must have an element $g \in P \setminus P'$. But, the first part of the lemma implies that $g \notin \mathrm{N}_G(P')$, i.e. $gP'g^{-1}$ is a maximal p–subgroup of G distinct from P'. That is $P' \notin \left(\mathcal{Syl}_p(G)\right)^P$ as desired. ■

The most important part of this is the fact that

COROLLARY 13.31. *Let P be as before. Suppose that $\mathcal{X} \subseteq \mathcal{Syl}_p(G)$ is P–invariant under the action of conjugation [II]. Then*

$$\#\mathcal{X} \equiv_p 1$$

if $P \in \mathcal{X}$ and

$$\#\mathcal{X} \equiv_p 0$$

otherwise.

PROOF. We have just proven that there is only one global fixed point for this action, namely the group P itself. Therefore P is not in $\mathcal{X}$, then there are no global fixed points for the action of P on $\mathcal{X}$. On the other hand if P is in $\mathcal{X}$ then there is exactly 1, namely P itself. Therefore **OPST** gives the result. ■

[II] Written out, this means for all $g \in P$ and $X \in \mathcal{X}$ we have that $gXg^{-1} \in \mathcal{X}$)

13.4.2 Proofs of Sylow's Theorems

Proof. (**Proof of Sylow1**) We prove that

$$\text{every } p\text{–subgroup of } G \text{ is contained in a Sylow } p\text{–subgroup of } G. \tag{2}$$

By Cauchy's lemma, there exists an element of order p in G, and as always, this element generates a subgroup of G which has p elements. Therefore, proving (2) will prove **Sylow1**.

Let P be a p–subgroup of G which is not properly contained inside of any other p–subgroup of G. We prove (2) by proving that P is a Sylow p–subgroup of G. By way of contradiction we assume that this is false, that is we assume P is not a Sylow p–subgroup of G. This means that p divides the index of P in G, $\#(G/P)$.

Since

$$\begin{aligned}
\#\left(G/P\right) &= \frac{|G|}{|P|} \\
&= \frac{|G|}{|P|} \cdot \frac{|\mathrm{N}_G(P)|}{|\mathrm{N}_G(P)|} \\
&= \frac{|G|}{|\mathrm{N}_G(P)|} \cdot \frac{|\mathrm{N}_G(P)|}{|P|} \\
&= \#\left(G/\mathrm{N}_G(P)\right) \cdot \left|\mathrm{N}_G(P)/P\right|
\end{aligned}$$

and since p is a prime number, p must divide at least one of these two factors

$$\#\left(G/\mathrm{N}_G(P)\right) \quad \text{or} \quad \left|\mathrm{N}_G(P)/P\right|.$$

We will prove that p divides neither, which will give us our contradiction.

By **OST** theorem we know that

$$\#(G/\mathrm{N}_G(P)) = \#\mathcal{X}$$

where $\mathcal{X}$ is defined (here) as the set of conjugate subgroups of P in G. This is because $\mathcal{X}$ is the orbit of $P \in \mathcal{X}$ under the action G on the set of subgroups of G, and the stabilizer of P under this action is $\mathrm{N}_G(P)$. Of course $P \in \mathcal{X}$ since any subgroup is conjugate to itself. Moreover, the restriction action of P on $\mathcal{S}\mathrm{yl}_p(G)$ preserves $\mathcal{X}$, and by Corollary 13.31, we therefore have that this number, $\#(G/\mathrm{N}_G(P))$ is equivalent to 1 mod p, in particular p doesn't that number.

Meanwhile, part 2 of Lemma 13.29 says that p also does not divide $|\mathrm{N}_G(P)/P|$. Hence there is no way that p could divide $\#(G/P)$ and hence P must be a Sylow p–subgroup. ∎

Remark 13.32. What we really just proved is that every p–subgroup of G is contained in a Sylow p–subgroup of G.

Proof. (**Proof of Sylow2**) Let P and P' be Sylow p–subgroups of G which are not conjugate and let

$$\mathcal{X} := \{gP'g^{-1} : g \in G\}.$$

That is $\mathcal{X}$ is the set of all subgroup of G which are conjugate to P'. Since P is not conjugate to P' by assumption, $P \notin \mathcal{X}$. However taking $g = e_G$, P' clearly is (every subgroup is conjugate to itself). Therefore by Corollary 13.31 we must have that the number of element of $\mathcal{X}$ is both divisible and not divisible (by taking $P' = P$ in statement of Corollary) by p, a contradiction. ∎

Proof. (**Proof of Sylow3**) Another way of stating **Sylow2**, which is what we just proved, is that the action $G \circlearrowright \mathbf{Syl}_p(G)$ given by conjugation is transitive. Therefore, **OST** implies that

$$n_p(G) = \#\left(\mathbf{Syl}_p(G)\right) = \#\left({}^{G}\!/_{\mathrm{N}_G(P)}\right)$$

where P is any Sylow p–subgroup of G. Therefore we must have that

$$n_p(G) \cdot |\,\mathrm{N}_G(P)\,| = |\,G\,|\,.$$

In particular, $n_p(G) \Big| |\,G\,|$. On the other hand, for a Sylow p–subgroup P we have that $P \circlearrowright \mathbf{Syl}_p(G)$ and $P \in \mathbf{Syl}_p(G)$. Therefore by Corollary 13.31 we see that the number of elements of $\mathbf{Syl}_p(G)$ – which we have been denoting by $n_p(G)$ – is equivalent to 1 mod p, as desired. ■

13.5 More Examples/Exercises

Example 13.33. Let's consider the group $G = \mathbf{SL}_2(\mathbb{Z}_7)$. We have seen already in Example/Theorem 10.98, that this group group has

$$8 \cdot 7 \cdot 6 = 336 = 7 \cdot 3 \cdot 2^4$$

elements. Now if we consider the subgroup G generated by the matrices

$$\mathbf{A} = \begin{pmatrix} 2 & 5 \\ 2 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$$

we can see easily that $|\,\mathbf{A}\,| = 8$ (since the trace is a square root of 2), $|\,\mathbf{B}\,| = 4$(since the trace is 0) and their product

$$\mathbf{A} \cdot \mathbf{B} = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}$$

which also has order 4 since it's trace is 0. Hence

$$(\mathbf{AB})^2 = -\mathbf{I}_2 \Rightarrow \mathbf{BAB} = -\mathbf{A}^{-1} \Rightarrow \mathbf{BA}(-\mathbf{B}) = \mathbf{A}^{-1} \Rightarrow \mathbf{BAB}^{-1} = \mathbf{A}^{-1}$$

where we have used the fact that a matrix has order 4 and determinant 1 iff it's square is $-\mathbf{I}_2$ several times. Lastly note that $\mathbf{A}^4 = \mathbf{B}^2$ since both have order 2 and therefore equal to $-\mathbf{I}_2$, hence each other. These facts imply that $\langle\, \mathbf{A}, \mathbf{B} \,\rangle$ has order 16 and is isomorphic to $B\mathcal{D}_4$. In particular it is the Sylow 2-subgroup of $\mathbf{SL}_2(\mathbb{Z}_7)$. Since the order of the Sylow 7 and Sylow 3–subgroups are each prime, they are isomorphic to $\mathbb{Z}_7$ and $\mathbb{Z}_3$ respectively.

Exercise 13.14. This exercise is a lemma for the next one. Show for all positive prime numbers p, that there is a solution to the equation

$$a^2 + b^2 = -1.$$

SOLUTION

Hint! If there exists an element $x \in \mathbb{Z}_p$ with $x^2 = -1$, then we're done. Simply take $a = x$ and $b = 0$. Otherwise, if not, then show for every perfect square y that $-y$ is not a perfect square and vice versa. Then argue as follows: 1 is always a perfect square and set b s.t. $b^2 = 1$. Then we can find an a that works iff -2 is a perfect square, i.e. iff 2 is not a perfect square. if 2 is a perfect square then let b be a square root of 2 and then the equation has a solution iff -3 is a perfect square. That is iff 3 is not a perfect square. Thus we are still screwed iff 3 is also a perfect square. Continue like this to show that all elements of $\mathcal{U}(\mathbb{Z}_p)$ are perfect squares, which is nonsense[1], since we have shown that half of all non-zero elements are perfect squares.

EXERCISE 13.15. Let p be an odd prime and suppose that $p^2 - 1$ is not divisible by 16. (It's easy to see that it's always divisible by 8 since both $p - 1$ and $p + 1$ are even and at least one of them must be divisible by 4 since every other even number is divisible by 4.) Prove that for all odd primes that $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ contains an isomorphic copy of Q_8. Conclude that if $p^2 - 1$ is not divisible by 16[II] that the Sylow 2-subgroup of $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ has Q_8 as its isomorphism type.

SOLUTION

Hint! Let $\mathbf{J}$ be the matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

which has order 4. Suppose that $\mathbf{A}$ is a (another) matrix of order 4, determinant 1, and satisfies the relation

$$(\mathbf{J} \cdot \mathbf{A})^2 = -\mathbf{I}_2.$$

Show that $\langle\, \mathbf{J}, \mathbf{A} \,\rangle$ is the Sylow 2–subgroup of $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ and is isomorphic to Q_8. Lastly, show that such a matrix $\mathbf{A}$ always exists. (You will come down to needing to show that $a^2 + b^2 = -1$ always has a solution in $\mathbb{Z}_p$. This is the lemma you just proved in the previous exercise!)

For the following exercise, you will need to use the following result, proven in the appendix, (there it's Theorem G.10 coming up on page 842).

LEMMA 13.34. *Suppose $u \in \mathbb{Z}_p$ is not a perfect square. Then, the set of non–zero matrices of the form*

$$\begin{pmatrix} a & bu \\ b & a \end{pmatrix}$$

form a cyclic subgroup of $\mathbf{GL}_2\left(\mathbb{Z}_p\right)$ of order $p^2 - 1$. ■

The fact that this is an abelian group of order $p^2 - 1$ is straight forward. (To show that all of these matrices have non–zero determinant uses the fact that u is assumed to not be a perfect square.) The only hard part is proving that this group is indeed cyclic. The proof involves showing that this set of matrices are the set of invertible elements in some finite field.

EXERCISE 13.16. (Quadratic Reciprocity) Prove that $[\,2\,]_p$ is a perfect square in $\mathbb{Z}_p$ (actually in $\mathcal{U}(\mathbb{Z}_p)$) iff 16 divides $p^2 - 1$.

SOLUTION

Hint! We have previously shown that $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ contains an element of order 8 iff we can solve the stated equation. Now, the previous exercise showed that the Sylow 2–subgroup of $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ is Q_8 whenever 16 does not divide $p^2 - 1$. In particular, such a group does not contain an element of order 8.

It remains to prove the converse, if 16 divides $p^2 - 1$, then $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ contains an element of order 8, which can be a bit of a challenge without using the above result from the appendix. There are two (mutually exclusive) cases.

If $4 \mid (p - 1)$, then 4 cannot divide $p + 1$, so $8 \mid (p - 1)$. Prove this implies that $\mathcal{U}(\mathbb{Z}_p)$ contains an element of order 8, ζ. Use this element to get the matrix

$$\begin{pmatrix} \zeta & 0 \\ 0 & \zeta^{-1} \end{pmatrix}$$

which has determinant 1 and order 8.

[I] This gives an example of an equation which is solvable in $\mathbb{Z}_p$ for every prime, but which has no solution in the integers (or even the real numbers!) There are even polynomials with a single variable for which this is the case!

[II] The first few primes fitting into this category are 7, 17, 23, 31,33, 39, 41.

If 4 does not divide $p-1$, then $\mathcal{U}(\mathbb{Z}_p)$ does not contain an i. Moreover, the subgroup in the lemma being cyclic, contains an element (i.e. a matrix) $\mathbf{A}$ of order 16. The only problem with $\mathbf{A}$ is that it may not have determinant 1. But, its not hard to show that (show it!) the assumption that $\mathcal{U}(\mathbb{Z}_p)$ does not contain an i implies that $\mathbf{A}^2$ does have determinant 1. Therefore $\mathbf{A}^2$ is our element of order 8 in this case.

Exercise 13.17. Show that $\mathbf{SL}_3(\mathbb{Z}_2)$ has order table

Order	1	2	3	4	7
Number	1	21	56	42	48

SOLUTION

(In other words, the group $\mathbf{SL}_3(\mathbb{Z}_2)$ contains 1 element of order 1, 21 elements of order 2, 56 elements of order 3, 42 elements of order 4, and 48 elements of order 7).
Hint! There must be 8 Sylow 7–subgroups and 28 Sylow 3–subgroups (why?? Calculate normalizers!). Therefore a total of 48 elements of order 7 and 56 elements of order 3. To figure out the number of elements of order 2 is a bit trickier, (since the Sylow 2–subgroup is $\cong \mathcal{D}_4$ and is not cyclic.) Now the (a) Sylow 2–subgroup of $\mathbf{SL}_3(\mathbb{Z}_2)$ can be described as $\mathrm{UT}_3(\mathbb{Z}_2)$. Indeed, we have

$$|\mathbf{SL}_3(\mathbb{Z}_2)| = 168 = 2^3 \cdot 3 \cdot 7 \quad \text{and} \quad |\mathrm{UT}_3(\mathbb{Z}_2)| = 2^3.$$

Next, check that the centralizer of the element

$$z := \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

consists only of $\mathrm{UT}_3(\mathbb{Z}_2)$.

We've seen a few times already that $\mathrm{UT}_3(\mathbb{Z}_2) \cong \mathcal{D}_4$ generated by the two elements of order 2,

$$x := \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad y := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

The isomorphism $\mathcal{D}_4 \to \mathrm{UT}_3(\mathbb{Z}_2)$ takes $\rho \mapsto xy$ and $r \mapsto x$. Thinking in $\mathcal{D}_4$ we see that there are 3 conjugacy classes of elements of order 2 which are represented by x, y, and z. Lastly, check in the bigger group that these three matrices are conjugate. Conclude that all elements of order 2 are conjugate in $\mathbf{SL}_3(\mathbb{Z}_2)$. Then use your calculation of the size of the centralizer to conclude that there are $168/8 = 21$ elements of order 2 in $\mathbf{SL}_3(\mathbb{Z}_2)$.

As for elements of order 4, show that all elements of order 4 are conjugate to $x \cdot y$ and that the centralizer of this matrix, $x \cdot y$, has order 4. (To make your life easier, note if a matrix centralizes $x \cdot y$ it must centralize it's square, which is z.) By doing an addition verify that all 168 elements of $\mathbf{SL}_3(\mathbb{Z}_2)$ are now accounted for.

Exercise 13.18. What is the smallest odd integer which can be the order of a non-abelian group. After the one you just found, what is the second smallest such integer?
Hint! It's not 15.

SOLUTION

Exercise 13.19. Under the action of $\mathcal{S}_6$ on its 10 Sylow 3–subgroups (first discussed in Example 13.19 back on page 688), explicitly write down what the cycle type of the image (under the permutation representation) of a

SOLUTION

(1) 2–cycle
(2) 2 2–cycle
(3) 2 2 2–cycle
(4) 3–cycle
(5) 3 3–cycle
(6) 4–cycle
(7) 4 2–cycle
(8) 5–cycle
(9) 6–cycle

is.

Hint! You can do this by finding where your favorite generating set ends up, and then take products. Or if want, you could do this case by case. For example, a 2–cycle normalizes only the 4 Sylow 3–subgroups which contain

$$(1\,2\,3) \quad (1\,2\,4) \quad (1\,2\,5) \quad (1\,2\,6).$$ [III]

Moreover, since a 2–cycle has image order 2 in $\mathcal{S}_{10}$, it must be a product of 2–cycles which we now know fixes 4 letters (the four letters fixing the above 4 indicated Sylow 3–subgroups.) Therefore a 2–cycle must become a 2 2 2–cycle.

13.6 Fundamental Theorem of Finite Abelian Groups

Of course every cyclic group is a abelian, and every direct product of cyclic groups is also abelian. The Fundamental Theorem of Finite Abelian Groups says that this covers all abelian groups.

Theorem 13.35 (Fundamental Theorem of Finite Abelian Groups)**.** *Every finite abelian group is the direct product of cyclic groups.*

It is the goal of this section to prove this theorem.
Let's outline our plan of attack:

Plan of Attack 13.36.

Step 1 Prove that every finite abelian group is a direct product of Abelian p-groups. This reduces the problem of classifying abelian groups to the problem of classifying abelian p-groups.

Step 2 Prove that every finite abelian p-group is isomorphic to some $\mathbb{Z}_{p^\lambda}$ (defined below) for some partition $\boldsymbol{\lambda}$.

Step 3 Prove that the partition uniquely determines the abelian p-group, $\mathbb{Z}_{p^\lambda}$. In other words $\mathbb{Z}_{p^\lambda} \cong \mathbb{Z}_{p^{\lambda'}}$ implies that the two partitions $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}'$ are the same.

Step 4 Conclude from the previous two steps that partitions of n and abelian group of order p^n are in bijection: Abelian groups of order $p^n \overset{1\text{-}1}{\leftrightsquigarrow}$ Partitions of n

Proposition 13.37. *If n is factored as $p_1^{n_1} \cdots p_k^{n_k}$ according to the FTOA into distinct prime powers, then any abelian group of order n is isomorphic to*

$$P_1 \times \cdots \times P_k$$

[III] These are (resp.)

$$\langle (1\,2\,3),(4\,5\,6)\rangle, \quad \langle (1\,2\,4),(3\,5\,6)\rangle, \quad \langle (1\,2\,5),(3\,5\,6)\rangle, \quad \text{and} \quad \langle (1\,2\,6),(4\,5\,3)\rangle.$$

where P_i is an abelian group of order $p_i^{n_i}$. Moreover,

$$P_1 \times \cdots \times P_k \cong P_1' \times \cdots \times P_k'$$

iff $P_i \cong P_i'$ for all $1 \leq i \leq k$.

Proof. Let G be an abelian group of order n. Take each P_i to be a Sylow p_i–subgroup of G. Then, like any subgroup of an abelian group, each $P_i \trianglelefteq G$. But this implies that each Sylow subgroup of G is normal in G, so

$$G \cong P_1 \times \cdots \times P_k.$$

Moreover, since the P_i's are the Sylow p–subgroups of G, we can apply Theorem 13.25 to be able to conclude the second statement. ■

Therefore, we simply need to classify abelian p–groups, i.e. abelian groups whose order is a power of a prime.

Example 13.38. Let's look at some obvious abelian groups of order $3^5 = 243$. It will turn out in general, that the only abelian groups are the obvious ones which are

$$\mathbb{Z}_{3^5} \qquad \mathbb{Z}_{3^4} \times \mathbb{Z}_3 \qquad \mathbb{Z}_{3^3} \times \mathbb{Z}_{3^2}$$
$$\mathbb{Z}_{3^3} \times \mathbb{Z}_3 \times \mathbb{Z}_3 \qquad \mathbb{Z}_{3^2} \times \mathbb{Z}_{3^2} \times \mathbb{Z}_3 \qquad \mathbb{Z}_{3^2} \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$$
$$\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3.$$

Notice that, just checking out the exponents, we always get a partition of the number 5. Moreover, all 7 partitions of 5 show up this way. It will also follow from the general theory, or you can check it directly yourself, that no two of the 7 groups listed here are isomorphic. There is nothing special about the number 2 here and we could just as well substitute all $3's$ above with 2's (to give us examples of abelian groups of order $2^5 = 32$) or all $5's$ to give us abelian groups of order $5^5 = 3125$.

Recall that for any natural number n, an ordered k–tuple of positive integers $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_k)$ is a partition of n if

(1) $\lambda_1 \geq \cdots \geq \lambda_k$

(2) $\lambda_1 + \cdots + \lambda_k = n$.

Of course, $k \leq n$. To limit special cases, we can define $\lambda_i = 0$ for $k < i \leq n$. In this way we can think of any partition of consisting of n terms. Of course with our original requirement, only the partition consisting of n 1's, $(1, \ldots, 1)$ has exactly n terms. The other extreme is the partition of n which consists of a single (non-zero) term (n). Further recall that $\pi(n)$ denotes the number of partitions of n. For convience, we rerecord the table from page 365 where we first discussed partitions and how to calculate $\pi(n)$.

Definition 13.39. For a partition of $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_k)$, and a prime number p, define the group

$$\mathbb{Z}_{p^{\boldsymbol{\lambda}}} = \mathbb{Z}_{p^{\lambda_1}} \times \cdots \times \mathbb{Z}_{p^{\lambda_k}}.$$

Of course this is an abelian group of order $p^{\lfloor \boldsymbol{\lambda} \rfloor}$.

Example 13.40.

$$\begin{aligned} \mathbb{Z}_{3^{(4,4,3,2,2,1,1)}} &= \mathbb{Z}_{3^4} \times \mathbb{Z}_{3^4} \times \mathbb{Z}_{3^3} \times \mathbb{Z}_{3^2} \times \mathbb{Z}_{3^2} \times \mathbb{Z}_{3^2} \times \mathbb{Z}_{3^1} \times \mathbb{Z}_{3^1} \\ &= \mathbb{Z}_{81} \times \mathbb{Z}_{81} \times \mathbb{Z}_{27} \times \mathbb{Z}_9 \times \mathbb{Z}_9 \times \mathbb{Z}_9 \times \mathbb{Z}_3 \times \mathbb{Z}_3. \end{aligned}$$

n	1	2	3	4	5	6	7	8	9	10
$\pi(n)$	1	2	3	5	7	11	15	22	30	42
n	11	12	13	14	15	16	17	18	19	20
$\pi(n)$	56	77	101	135	176	231	297	385	490	627
n	21	22	23	24	25	26	27	28	29	30
$\pi(n)$	792	1002	1255	1575	1958	2436	3010	2718	4565	5604
n	31	32	33	34	35	36	37	38	39	40
$\pi(n)$	6842	8349	10143	12310	14883	17977	21637	26015	31185	37338
n	41	42	43	44	45	46	47	48	49	50
$\pi(n)$	44583	53174	63261	75175	89134	105558	124754	147273	173525	204226

Table 13.1: Partitions of n

EXAMPLE 13.41. The abelian group of order $3^5 = 243$ in our new notation are

$$\mathbb{Z}_{3^{(5)}} = \mathbb{Z}_{3^5} \qquad \mathbb{Z}_{3^{(4,1)}} = \mathbb{Z}_{3^4} \times \mathbb{Z}_3 \qquad \mathbb{Z}_{3^{(3,2)}} = \mathbb{Z}_{3^3} \times \mathbb{Z}_{3^2}$$
$$\mathbb{Z}_{3^{(3,1,1)}} = \mathbb{Z}_{3^3} \times \mathbb{Z}_3 \times \mathbb{Z}_3 \qquad \mathbb{Z}_{3^{(2,2,1)}} = \mathbb{Z}_{3^2} \times \mathbb{Z}_{3^2} \times \mathbb{Z}_3 \qquad \mathbb{Z}_{3^{(2,1,1,1,1)}} = \mathbb{Z}_{3^2} \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$$
$$\mathbb{Z}_{3^{(1,1,1,1,1)}} = \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3.$$

13.6.1 EXISTENCE

We now move on with our plan by proving an existence theorem: namely we prove that for every abelian p-group there exists an partition λ (of some number) so that the given abelian group is isomorphic to $\mathbb{Z}_{p^\lambda}$. Later, we prove the converse, that it is isomorphic to only one of these dudes.

Since every finite abelian group is the direct product of its Sylow p–subgroups, each of which are abelian p–groups, we simply need to classify abelian p–groups. We write all abelian groups in this part of the proof additively. Fix a finite abelian group G. For $S = \{s_1, \ldots, s_m\}$ a subset G. Define

$$\| S \| = | s_1 | \cdot \cdots \cdot | s_m |$$

called the *length of* S. Of course if S' is a proper subset of S, then $\|S'\| < \|S\|$. Note that for the p–group $\mathbb{Z}_{p^\lambda}$, where λ is a partition of the number n, and for S the obvious generating set of $\mathbb{Z}_{p^\lambda}$, $\|S\| = p^r$.

PROPOSITION 13.42. *For every abelian p–group, G, of order p^n there exists a partition of n, λ, so that*

$$G \cong \mathbb{Z}_{p^\lambda}.$$

In particular, every abelian p–group is (isomorphic to) a direct product of cyclic p–subgroups.

Proof. Let $S := \{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$ be a generating set for G with minimal length. That is

$$\forall \widehat{S} \subseteq G \; \|S\| \leq \|\widehat{S}\| \quad \text{or} \quad \left\langle \widehat{S} \right\rangle \neq G \tag{3}$$

For one thing, this means that no proper subset of S generates G, and so no $\mathbf{s}_i$ can be in $\langle S \setminus \{s_i\} \rangle$ for then $\langle S \setminus \{\mathbf{s}_i\} \rangle$ would equal all of G.

Since G is a p–group, the order of every element of G, and hence the order of every element of S, is a power of the prime p, i,e, $|\mathbf{s}_i| = p^{\lambda_i}$ for some $\lambda_i > 1$. Now, from Example/Theorem 9.36 we have an epimorphism

$$\begin{aligned} \varphi &: \mathbb{Z}_{p^{\boldsymbol{\lambda}}} \to G \\ &: ([a_1]_{p^{\lambda_1}}, \ldots, [a_n]_{p^{\lambda_n}}) \mapsto a_1\mathbf{s}_1 + \cdots + a_n\mathbf{s}_n \end{aligned}$$

where $\boldsymbol{\lambda}$ is the partition (after reordering the λ_i so they are in non-decreasing order) $(\lambda_1, \ldots, \lambda_n)$. If φ is injective, then it's an isomorphism and we are done. Therefore, we assume that φ is not injective. In other words there exists a non-trivial element

$$\exists([a_1]_{p^{\lambda_1}}, \ldots, [a_n]_{p^{\lambda_n}}) \in \operatorname{Ker}(\varphi), \; ([a_1]_{p^{\lambda_1}}, \ldots, [a_n]_{p^{\lambda_n}}) \neq (0, \ldots, 0). \tag{4}$$

So we have

$$a_1\mathbf{s}_1 + \cdots + a_n\mathbf{s}_n = 0$$

where we can assume that each $0 \leq a_i < p^{\lambda_i}$ and at least some a is non-zero. Let $r \geq 0$ be the smallest integer so that $p^r | a_i$ for each a_i. This means there is an m so that $p^r | a_m$ but p^{r+1} does not divide a_m.

Define the integers $b_i := a_i / p^r$. We have

$$p^r (b_1\mathbf{s}_1 + \cdots + b_n\mathbf{s}_n) = 0 \tag{5}$$

Note that b_m is now relatively prime with p. Moreover, by multiplying both sides of (5) by a multiplicative inverse of b_m in $\mathbb{Z}_{p^{\lambda_m}}$, we can assume that $b_m = 1$. Now let's define $\mathbf{s}$ as

$$\mathbf{s} := b_1\mathbf{s}_1 + \cdots + b_n\mathbf{s}_n \in G.$$

By (5), $p^r \cdot \mathbf{s} = 0$, so we know that $|\mathbf{s}| \leq p^r$.

Let's start putting this together. We just explained that $|\mathbf{s}| \leq p^r$. Since $p^r | a_m$, we know that $p^r \leq a_m$. Next, our assumption on each a_i is that $a_i < p^{\lambda_i} = |\mathbf{s}_i|$. In particular, $a_m < |\mathbf{s}_m|$. Connecting the dots, we have that $|\mathbf{s}| < |\mathbf{s}_m|$. Moreover since the coefficient $b_m = 1$, we can use the defining equation of $\mathbf{s}$ to solve for $\mathbf{s}_m$ in terms of $\mathbf{s}_j$, $j \neq m$ and $\mathbf{s}$. This means if we define S' as the set we get from replacing $\mathbf{s}_m$ with $\mathbf{s}$ in S, i.e.

$$S' := (S \setminus \mathbf{s}_m) \cup \{\mathbf{s}\},$$

then S' also generates G. Lastly, since $|\mathbf{s}| < |\mathbf{s}_m|$, we also have that $\|S'\| < \|S\|$. Together, these last two facts contradict (3). Therefore (4) must be false and φ is an isomorphism. ∎

Remark 13.43. We actually proved that if we take a generating set $\{s_1, \ldots, s_m\}$ of G of smallest length, then

$$G \cong \mathbb{Z}_{p^{\boldsymbol{\lambda}}}$$

where $\mathbb{Z}_{p^{\boldsymbol{\lambda}}}$ is the partition consisting of λ_i where $|s_i| = p^{\lambda_i}$.

13.6.2 Uniqueness

We do much more in this subsection than prove uniqueness. We actually figure out the number of elements of a given order in an abelian p–group based on the associated partition. This is best done in terms of the dual of the given partition.

Definition 13.44. For a partition λ of n, we define the *dual partition* of λ, $\lambda^\top$, as

$$\mu = (\mu_1, \ldots, \mu_m)$$

where

$$\mu_i = \#\{j \in \overline{k} : \lambda_j \geq i\}.$$

For instance, if λ has k (non-zero) terms, we always have that $\mu_1 = k$.

Example 13.45. We have $(1, \ldots, 1)^\top = (n)$ since there are n terms greater than or equal to 1, but none greater than or equal to 2. We also have that $(1, \ldots, 1) = (n)^\top$ since the first and only term of (n) gives us one and only term greater than 1, one and only term greater than 2, etc.

To prove the following lemma, we give a way to visualize a partition and its dual. For any partition,

$$\lambda = (\lambda_1, \ldots, \lambda_k)$$

of n, the *Young diagram* of λ is simply an array with λ_i boxes in the i^{th} row of the array colored in. For instance, the partition $(4, 4, 3, 2, 2, 1)$ of

$$16 \quad (= 4 + 4 + 3 + 2 + 2 + 1)$$

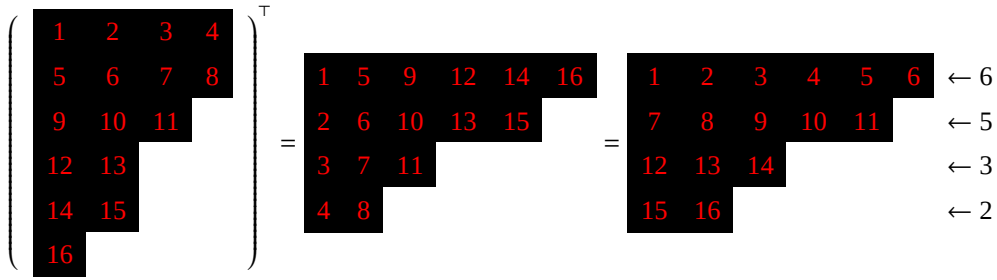
has Young diagram.

1	2	3	4	← 4
5	6	7	8	← 4
9	10	11		← 3
12	13			← 2
14	15			← 2
16				← 1

.

(The numbers written are not part of the definition, they are simply there to help keep track of locations. The number at the end of each row simply is the number of boxes in the given row.) Now if we switch the rows and columns of the Young diagram of a partition, we end up with the Young diagram of the dual partition. This is because if there are i entries in the j^{th} column of a diagram, then there must be exactly i elements of the partition who are at least j. For instance, in our example, the dual partition has

Young diagram



which allows us to visualize the fact that $(4,4,3,2,2,1)^\top = (6,5,3,2)$. Note that what results is still a partition of 16. This should be clear since you do not change the number of boxes when you flip the diagram.

EXERCISE 13.20. Let

$$\lambda = (5,5,3,3,3,1,1),$$

which is a partition of 21. Compute $\lambda^\top$ and verify that it is also a partition of 21.

SOLUTION

LEMMA 13.46. *For any partition, λ, of n,*

(1) *The dual of λ, $\lambda^\top = \mu$, is also a partition of this same integer n.*

(2) *Transposition is an involution, that is $(\lambda^\top)^\top = \lambda$.*

PROOF.

(1) Clearly transposing the Young diagram of a partition does not change the number of boxes in it. Thus the sum of the entires of μ also equals n. Next, it's easy to see, even directly from the definition, that the entries of μ are decreasing. This is because there cannot be more entries of λ greater than a larger number than the number of entries greater than a smaller (duh!). Therefore, by definition, μ is a partition of m.

(2) If we switch the rows and columns of any array twice we have done nothing. ∎

COROLLARY 13.47. *Two partitions are equal iff their duals are equal.*

PROOF.

$$\lambda^\top = \lambda'^\top \Rightarrow \lambda = (\lambda^\top)^\top = (\lambda'^\top)^\top = \lambda'.$$ ∎

Recall that for a group G and a positive integer m, we defined the numbers

$$\mathbf{o}_m(G) \stackrel{\text{def}}{=} \#\{g \in G : |g| = m\} \quad \text{and} \quad \mathbf{O}_m(G) \stackrel{\text{def}}{=} \#\{g \in G : g^m = e_G\}.$$

Since g has order m iff $g^m = e_G$, and $g^{m/p} \neq e_G$ for every prime divisor of m. Since p is the only prime divisor of p^i, we see

$$\mathbf{o}_{p^i}(G) = \mathbf{O}_{p^i}(G) - \mathbf{O}_{p^{i-1}}(G). \tag{6}$$

THEOREM 13.48.

$$\mathbf{O}_{p^i}(\mathbb{Z}_{p^\lambda}) = p^{\mu_1+\cdots+\mu_i} \quad \textit{and} \quad \mathbf{o}_{p^i}(\mathbb{Z}_{p^\lambda}) = p^{\mu_1+\cdots+\mu_i} - p^{\mu_1+\cdots+\mu_{i-1}}.$$

where $\mu = \lambda^\top$.

Proof. Since

$$\mathbf{O}_{p^i} = \mathbf{o}_{p^i} - \mathbf{o}_{p^{i-1}}$$

by (6), we only need to prove the first formula. We do this by proving that

$$\frac{\mathbf{O}_{p^i}(\mathbb{Z}_{p^\lambda})}{\mathbf{O}_{p^{i-1}}(\mathbb{Z}_{p^\lambda})} = p^{\mu_i} \tag{7}$$

and using induction.

We've seen before the obvious fact that $\mathbf{O}_m$ is multiplicative in the sense that

$$\mathbf{O}_m(G_1 \times \cdots \times G_k) = \mathbf{O}_m(G_1) \cdot \ \cdots \ \cdot \mathbf{O}_m(G_k)$$

for any groups $G_1, \ldots, G_k$ and integer m.

Also, recall that for the cyclic group, $\mathbb{Z}_m$, of order m $\mathbf{O}_i(\mathbb{Z}_m) = \gcd(i, m)$. Moreover $\gcd(p^i, p^j) = p^{\min(i,j)}$. Therefore

$$\mathbf{O}_{p^i}(\mathbb{Z}_{p^{\lambda_j}}) = p^{\min(i,\lambda_j)}.$$

So

$$\frac{\mathbf{O}_{p^i}(\mathbb{Z}_{p^\lambda})}{\mathbf{O}_{p^{i-1}}(\mathbb{Z}_{p^\lambda})} = \prod_{j=1}^{k} \left(\frac{p^{\min(i,\lambda_j)}}{p^{\min(i-1,\lambda_j)}} \right). \tag{8}$$

Lastly, note that

(1) $p^{\min(i,\lambda_j)} > p^{\min(i-1,\lambda_j)}$;

(2) $p^{\min(i,\lambda_j)} \neq p^{\min(i-1,\lambda_j)}$;

(3) $\lambda_j \geq i$;

(4) $p^{\min(i,\lambda_j)} = p^i$;

(5) $p^{\min(i,\lambda_j)} = p^i$ and $p^{\min(i-1,\lambda_j)} = p^{i-1}$;

(6) $\dfrac{p^{\min(i,\lambda_j)}}{p^{\min(i-1,\lambda_j)}} = p$

are clearly all equivalent.

Using these equivalences we see the product on the right side of (8) picks adds a p to the product exactly for the λ_j's for which $\lambda_j \geq i$ (and only contribute 1's otherwise). Therefore, we see that the right hand side of that equation

$$\prod_{j=1}^{k} \left(\frac{p^{\min(i,\lambda_j)}}{p^{\min(i-1,\lambda_j)}} \right) = \prod_{\{j:\lambda_j \geq i\}} p = p^{\mu_i}, \tag{9}$$

since the set set $\{j : \lambda_j \geq i\}$, by its very definition, has μ_i elements. Connecting (8) with (9) proves (7), which was our goal. Since, like for any group, $\mathbf{O}_1(\mathbb{Z}_{p^\lambda}) = 1$, we can use formula (7) to prove the theorem by induction.[I] ∎

[I] For instance, (7) says that

$$p^{\mu_1} = \frac{\mathbf{O}_p(\mathbb{Z}_{p^\lambda})}{\mathbf{O}_1(\mathbb{Z}_{p^\lambda})} \overset{1}{=} \mathbf{O}_p(\mathbb{Z}_{p^\lambda}) \quad \text{and} \quad p^{\mu_2} = \frac{\mathbf{O}_{p^2}(\mathbb{Z}_{p^\lambda})}{\mathbf{O}_p(\mathbb{Z}_{p^\lambda})} \overset{p^{\mu_1}}{\Rightarrow} p^{\mu_1+\mu_2} = \mathbf{O}_{p^2}(\mathbb{Z}_{p^\lambda})$$

and so on.

Corollary 13.49. *For partitions λ and λ'*

$$\lambda = \lambda' \Leftrightarrow \mathbb{Z}_{p^\lambda} \cong \mathbb{Z}_{p^{\lambda'}}.$$

Proof. The forward direction is obvious since if $\lambda = \lambda'$, then $\mathbb{Z}_{p^\lambda}$ and $\mathbb{Z}_{p^{\lambda'}}$ are exactly the same groups. For the other direction, let's denote the duals of λ and λ' as μ and μ' as we have been doing. The Theorem says that

$$\frac{\mathbf{O}_{p^i}(\mathbb{Z}_{p^\lambda})}{\mathbf{O}_{p^{i-1}}(\mathbb{Z}_{p^\lambda})} = p^{\mu_i}.$$

Since for any group G, and any integer m, the number $\mathbf{O}_n(G)$ is an isomorphism invariant, we have proven that for any number i,

$$p^{\mu_i} = p^{\mu'_i}$$

and so $\mu = \mu'$. But we've seen in Corollary 13.47 that this implies that the original partitions, λ and λ' are also equal, as desired. ■

13.6.3 Conclusions and Examples

Proposition 13.50. *For any abelian group G of order p^n there exists a unique partition*

$$\lambda \text{ of } n$$

so that

$$G \cong \mathbb{Z}_{p^\lambda}.$$ ■

Corollary 13.51. *Up to isomorphism, there are a total of $\pi(n)$ abelian groups of order p^n.* ■

Remark 13.52. This is kinda wild. It says that the number of abelian groups of order p^n is completely independent of the prime p and depends only on the power n. For instance, if you check out Table I.2 coming up on page 870, in the appendix, you can see that while the number of non–abelian groups of order p^n certainly depends on p and n, (there are even exceptions to this!) the table seems to imply that the number of abelian groups of that same order depends only on n.

We glue things together via Proposition 13.37. For the next theorem and its corollary, m is a positive integer with prime factorization

$$p_1^{n_1} \cdots p_k^{n_k}.$$

Theorem 13.53. *For any abelian group G of order m there exist unique partitions, one for every prime divisor of m,*

$$\begin{gathered}\lambda_1 \text{ of } n_1 \\ \lambda_2 \text{ of } n_2 \\ \vdots \\ \lambda_k \text{ of } n_k\end{gathered}$$

so that

$$G \cong \mathbb{Z}_{p_1^{\lambda_1}} \times \cdots \times \mathbb{Z}_{p_k^{\lambda_k}}.$$ ■

Corollary 13.54. *Up to isomorphism, there are a total of $\pi(n_1) \cdots \pi(n_k)$ abelian groups of order $p_1^{n_1} \cdots p_k^{n_k}$.* ■

Written in another way, we have that

COROLLARY[2] 13.55. *Number of Abelian Groups of order* $m = \prod_{p \text{ prime}} \pi(\nu_p(m))$ ■

EXAMPLE 13.56. Every abelian group of order 243 is is isomorphic to exactly one of the 7 abelian groups listed listed in Example 13.38 back on page 703.

EXAMPLE 13.57. There are a total of $7 \cdot 5 \cdot 3$ abelian groups of order $2^5 \cdot 3^4 \cdot 7^3 = 889,056$. ($\pi(5) = 7$, $\pi(4) = 5$, and $\pi(3) = 3$.)

MAPLE EXAMPLE 13.58. The following code gives us a command in MAPLE, called nuAb, which outputs the number of abelian groups of a given (finite) order.

```
>  FF:=n->nops(ifactors(n)[2]);
>  nuAb:=proc(n)
     local i, flag, hold;
     flag:=1;
     hold:=ifactors(n)[2];
     for i from 1 to FF(n)
     do  flag:=flag*(numbpart(hold[i][2]));
     od;
     return(flag);
   end;
```

For instance, when I type in

```
>  nuAb(342144);
```

into maple I get an output of 105. This represents the fact that, up to isomorphism, there are 105 abelian groups of order 342, 144. Indeed, 342, 144 has a prime factorization $2^7 \cdot 3^5 \cdot 11$, while $\pi(7) = 15$, $\pi(5) = 7$, and of course $\pi(1) = 1$. Taking this product, we get $105 = 15 \cdot 7 \cdot 1$.

EXAMPLE 13.59. Consider the group:

$$\mathbb{Z}_{5^4} \times \mathbb{Z}_{5^4} \times \mathbb{Z}_{5^2} \times \mathbb{Z}_5.$$

The partition of 11 is $11 = 4 + 4 + 2 + 1$, and the dual partition is $11 = 4 + 3 + 2 + 2$.

n	1	5	$5^2 = 25$	$5^3 = 125$	$5^4 = 625$
$\mathbf{o}_n$	1	$5^4 - 1$	$5^4 \cdot (5^3 - 1)$	$5^7 \cdot (5^2 - 1)$	$5^9 \cdot (5^2 - 1)$
=	1	624	77,500	1,875,000	46,8750,000

.

EXAMPLE 13.60. In the group, $\mathcal{U}(\mathbb{Z}_{16})$ we can easily see

$$1,\ 7,\ 9,\ 15$$

are the solutions to the equation $x^2 = 1$ over $\mathbb{Z}_{16}$, (so $\mathbf{O}_2(\mathcal{U}(\mathbb{Z}_{16})) = 4$) and, in addition, to these four,

$$3,\ 5,\ 11,\ 13$$

are the solutions to the equation $x^4 = 1$ over $\mathbb{Z}_{16}$, (so $\mathbf{O}_2(\mathcal{U}(\mathbb{Z}_{16})) = 8$). Therefore $\mu_1 = 2$ and $\mu_2 = 1$, and so $\lambda_1 = 2$ and $\lambda_2 = 1$. As such,

$$\mathcal{U}(\mathbb{Z}_{16}) \cong \mathbb{Z}_{2^2} \times \mathbb{Z}_2 = \mathbb{Z}_4 \times \mathbb{Z}_2.$$

Recall that for a finite cyclic group A we have that $\mathbf{O}_n(A) = \gcd(|A|, n)$. In particular, for all integers we have that

$$\mathbf{O}_n(A) \leq n$$

for all integers. The following theorem gives a converse to this fact.

THEOREM 13.61. *Let A be a finite abelian group. Then* TFAE

(1) *A is cyclic;*

(2) *For all prime numbers p, $\mathbf{O}_p(A) \leq p$;*

(3) *For all integers n, $\mathbf{O}_n(A) \leq n$.*

PROOF. Exercise. ∎

EXERCISE 13.21.

(1) List all abelian groups of order 2310.

(2) List all abelian groups of order 448.

(3) List all abelian groups of order 432.

SOLUTION

EXERCISE 13.22. Prove the theorem, Theorem 13.61.

SOLUTION

Recall that

$$|\mathcal{U}(\mathbb{Z}_{2^n})| = \vartheta(2^n) = 2^n \cdot \left(1 - \frac{1}{2}\right) = 2^{n-1}.$$

In particular they are 2-groups. We've seen previously, namely in Exercise 5.59 back on page 294, that for $n > 2$, these groups are never cyclic. We now show they are pretty close to being cyclic.

LEMMA 13.62. *For all $n \geq 0$,*

$$\mathrm{Ker}\left(\hat{}^2 : \mathcal{U}(\mathbb{Z}_{2^n}) \to \mathcal{U}(\mathbb{Z}_{2^n})\right) = \{1, 2^n - 1, 2^n + 1, -1\}.^{\text{II}}$$

PROOF. We can check this directly for $n \leq 2$. We now proceed by induction with $n \geq 3$. Note that under the assumption that $n \geq 3$ we have that

$$n \geq 3 \Rightarrow 2n \geq n + 3 \Rightarrow 2n - 2 \geq n + 1 \Rightarrow 2(n-1) \geq n + 1. \tag{10}$$

[II] This is the same as the set $\mathcal{U}(\mathbb{Z}_{2^n})[2]$. It is the set of elements that square to 1 mod 2^n.

Denote the homomorphism as follows

$$\begin{aligned}\pi := \pi_{2^n}^{2^{n+1}} : \mathcal{U}(\mathbb{Z}_{2^{n+1}}) &\to \mathcal{U}(\mathbb{Z}_{2^n}) \\ : [\,a\,]_{2^{n+1}} &\mapsto [\,a\,]_{2^n}.\end{aligned}$$

Note that π is two to one with $[\,a\,]_{2^{n+1}}$ and $[\,a+2^n\,]_{2^{n+1}}$ being the two elements which map to $[\,a\,]_{2^n}$ under π. If an element is in $\mathcal{U}(\mathbb{Z}_{2^{n+1}})[2]$ then its image under π is in $\mathcal{U}(\mathbb{Z}_{2^{n+1}})[2]$. Of course the converse is false, but it at least narrows down where we need to look to find elements of order 2. Indeed, let $[\,a\,]_{2^n} \in \mathcal{U}(\mathbb{Z}_{2^{n+1}})[2]$. By induction, this means that (its image under π) $[\,a\,]_{2^n}$ is either $[\,2^{n-1}-1\,]_{2^n}$, $[\,2^{n-1}+1\,]_{2^n}$, $[\,1\,]_{2^n}$, or $[\,2^n-1\,]_{2^n}$. I claim that anybody mapping to the first two under η do not have order 2 in the bigger group, $\mathcal{U}(\mathbb{Z}_{2^{n+1}})$. The 4 elements of $\mathcal{U}(\mathbb{Z}_{2^{n+1}})$ which map to elements of the first two types are

$$[\,2^{n-1}+1\,]_{2^{2n+1}} \qquad [\,2^{n-1}-1\,]_{2^{2n+1}} \qquad [\,-2^{n-1}-1\,]_{2^{2n+1}} \qquad [\,1-2^{n-1}\,]_{2^{2n+1}}.$$

The first two are written exactly the same way as their images, and the last two can easily be seen to have the same image as the first two by adding $2^n = 2\cdot 2^{n-1}$ to them.

Note that each of these elements can be represented in terms of the first, written as x, respectively as

$$x \qquad -x^{-1} \qquad -x \qquad x^{-1}. \tag{11}$$

Of course we only need to check the first of these, since you can visually see that the last two are the negatives of the first two. Indeed

$$\left(2^{n-1}-1\right)\cdot\left(2^{n-1}+1\right) = 2^{2(n-1)}-1 \equiv_{2^{n+1}} -1$$

where I've used Equation (10) to get rid of the power of 2. This means that the four elements in (11) all have the same order.[III]

So it remains to prove that x does not have order 2.

$$x^2 = (2^{n-1}+1)^2 = 2^{2(n-1)} + 2\cdot 2^{n-1} + 1 = 2^{2(n-1)} + 2^n + 1 \equiv_{2^{n+1}} = 2^n + 1$$

which does not equal 1. (Again, I've used Equation (10) to get rid of the power of 2). ∎

PROPOSITION 13.63. *For every positive integer n,*

$$\mathcal{U}(\mathbb{Z}_{2^n}) \cong \mathbb{Z}_2 \times \mathbb{Z}_{2^{n-2}}.$$

PROOF. WLOG we can assume that $n \geq 3$ since the lower cases case be checked directly. The group $\mathcal{U}(\mathbb{Z}_{2^n})$, having order 2^{n-1} is an abelian 2 group. Therefore there is some partition of $n-1$, $\lambda_1 \leq \cdots \leq \lambda_k$ so that

$$\mathcal{U}(\mathbb{Z}_{2^n}) \cong \mathbb{Z}_{2^{\lambda_1}} \times \cdots \times \mathbb{Z}_{2^{\lambda_k}}.$$

First note that

$$\mathcal{U}(\mathbb{Z}_{2^n})[2] = \{1, 2^{n-1}-1, 1+2^{n-1}, -1\},$$

and so $2 = \lambda_1^\top = k$. Therefore

$$\mathcal{U}(\mathbb{Z}_{2^n}) \cong \mathbb{Z}_{2^{\lambda_1}} \times \mathbb{Z}_{2^{\lambda_2}} = \mathbb{Z}_{2^m} \times \mathbb{Z}_{2^{(n-1)-m}}$$

with $m \leq (n-1)-m$. We are done if we can prove that $m = 1$.

Notice that $m = 1$ iff there are only 8 elements in $\mathcal{U}(\mathbb{Z}_{2^n})[4]$, i.e. iff there are only $8-4 = 4$ elements of order 4 in $\mathcal{U}(\mathbb{Z}_{2^n})$.

We have $\mathcal{U}(\mathbb{Z}_{2^n})^{\times 2}$ is

$$\{[\,1+8a\,]_{2^{2n}} : a \in \mathbb{Z}[2^{n-3}]\}.$$

[III] Of course in any group, an element and its inverse have the same order. The fact that x and $-x$ have the same order follows directly from the fact that they differ by a central (well, the group is abelian) element, namely -1, of order 2 in a 2-group. (Check that this always happens in this case!)

Indeed, a perfect square in $\mathcal{U}(\mathbb{Z}_{2^n})$ must be a perfect square when taken mod 8, but the only perfect square mod 8 is $[\,1\,]_{2^3} = [\,1\,]_8$.

In particular, the only element of order 2 which is a perfect square is $[\,1 + 2^{n-1}\,]_{2^n}$. Meanwhile, since the homomorphism $\hat{}^2$ has 4 elements in its kernel, there are four elements of $\mathcal{U}(\mathbb{Z}_{2^n})$ when squared is $[\,1 + 2^{n-1}\,]_{2^n}$, i.e. there are 4 elements in $\mathcal{U}(\mathbb{Z}_{2^n})$ when squared have order 2, i.e. there are 4 elements in $\mathcal{U}(\mathbb{Z}_{2^n})$ of order 4. ■

[III] By checking orders using **KIT** and using the previous lemma, namely that the map $\hat{}^2 : \mathcal{U}(\mathbb{Z}_{2^n}) \to \mathcal{U}(\mathbb{Z}_{2^n})$ has 4 elements in its kernel, we also see the opposite inclusion:

$$\mathcal{U}(\mathbb{Z}_{2^n})^{\times 2} = \{[\,1 + 8a\,]_{2^{2^n}} : a \in \mathbb{Z}[2^{n-3}]\}$$

also holds, but we will not need this opposite inclusion.

III Additional Topics in Group Theory

14 Simple and Transitive Groups

Recall that a group is simple if its only normal subgroups are the identity and the group itself.

Proposition 14.1. *A group is simple iff every non-trivial homomorphism is injective.*

Proof. If a group G contains a proper, normal subgroup N, then $G \to G/N$ is a non injective homomorphism (since it has kernel N). Conversely, since the kernel of a homomorphism is a normal subgroup, the kernel must be all of G, hence the homomorphism is trivial, or the kernel is trivial (i.e. contains only the identity) in which case the homomorphism is injective. ■

Remark 14.2. Of course a non-abelian simple group is perfect. This was pointed out in Theorem 11.54.

So far we have proven only one family of groups to be simple, namely the cyclic groups of prime order, $\mathbb{Z}_p$. In fact we proved in Example/Theorem 6.26 that $\mathbb{Z}_p$ contains no PNT subgroups at all, never mind any normal PNT subgroups. Alternatively, we can see this as a consequence of Lagrange's theorem, since a prime number p has no (positive) divisors except for 1 and p itself.

Example 14.3. For p a prime, the group $\mathbb{Z}_p$ is simple.

Since every subgroup of an abelian group is normal, taking the subgroup of a finite abelian group generated by an element of prime order shows that the only, finite simple abelian groups are of this form, $\mathbb{Z}_p$. We explained this in more detail in Example/Theorem 8.51 back on page 410. Therefore we have (after much work ... right) classified to abelian simple groups.

Theorem 14.4 (Same as 8.51)**.** *A finite abelian group is simple iff it is isomorphic to* $\mathbb{Z}_p$. ■

Amazingly, the classification of non–abelian simple groups has also been done. However, unlike the one line proof which classifies abelian simple groups, the classification of non–abelian simple groups is a bit harder. In fact, the proof of the theorem consists of tens of thousands of pages in several hundred journal articles written by about 100 authors, published mostly between 1955 and 2004. Like I said, just a bit harder ☺.

In this chapter, since the focus is mostly on permutation groups, we reintroduce our notation from Chapter 7 and write $\sigma[i]$, instead of $\sigma(i)$, for the image of the letter i under a permutation σ. We therefore focus only non–abelian simple groups. Let's recall the following definition from Chapter 10.

Definition 14.5. A subgroup Ω of $\mathcal{S}_n$ is *m–transitive* if for every sequence of distinct letters $1 \le a_1, \ldots, a_m \le n$ there exists a $\sigma \in \Omega$ with

$$\sigma[1] = a_1, \ldots, \sigma[m] = a_m.$$

We call an 1–transitive subgroup a *transitive subgroup*.

In that chapter we proved several things about such actions. For instance, we explained how this definition is (obviously) the same as the requirement that Ω acts transitively on $\overline{n}^{*m}$.

RANT 14.6. This is one of the few concepts which are not defined up to isomorphism! For instance $\langle (1\,2\,3)(4\,5)\rangle$ is not a transitive subgroup of $\mathcal{S}_6$ (or even $\mathcal{S}_5$) but $\langle (1\,2\,3\,4\,5\,6)\rangle$ is.

In the worst case scenerio, there is even the case of thinking of $\mathcal{S}_5 \preccurlyeq \mathcal{S}_6$ in the usual way (as permutations which fix the letter 6) or in the way detailed in 10.78 given by (on generators)

$$\begin{aligned}\varphi &: \mathcal{S}_5 \to \mathcal{S}_6\\ &: (1\,2) \mapsto (1\,2)(3\,4)(5\,6)\\ &: (1\,2\,3\,4\,5) \mapsto (2\,6\,5\,4\,3).\end{aligned}$$

where, among other things, $\underline{2}$–cycles in $\mathcal{S}_5$ become $\underline{2}\,\underline{2}\,\underline{2}$–cycles in $\mathcal{S}_6$. Under the first (standard) monomorphism the image (which we safely call $\mathcal{S}_5$) is not transitive (since everybody fixes 6) but under the second monomorphism, the image (which we call $\widetilde{\mathcal{S}_5}$)is transitive. To make matters worse there is an automorphism of $\mathcal{S}_6$ which takes $\mathcal{S}_5 \to \widetilde{\mathcal{S}_5}$. This is despite the fact that each of these groups are isomorphic to $\mathbb{Z}_6$. Therefore, when discussing

To avoid confusion with other notation, we write the orbit of a letter under the action of a subgroup of $\mathcal{S}_n$ in the following way:

DEFINITION 14.7. Suppose that $\Omega \preccurlyeq \mathcal{S}_n$. We denote the orbit of $i \in \overline{n}$ under the action of Ω as i/Ω.

PROPOSITION 14.8.

PROOF. Suppose that i and j are in the same N orbit. Then we must have that there exists an $h \in N$ so that $h[i] = j$. Therefore

$$(gh)[i] = g[(h[i])] = g[j].$$

But since $N \trianglelefteq G$, there exists an $h' \in H$ so that $gh = h'g$. Therefore, we must have that $h'g[i] = g[j]$ and therefore G must preserve the orbits of N. In other words, if $\sim_N$ denotes the orbit relation (i.e. the relation which relates two letters iff they are in the same N orbit)

$$i \sim_N j \Rightarrow g[i] \sim_N g[j].$$

This means that

$$\begin{aligned}\bullet &: G \circlearrowright \overline{n}/_N\\ &: g \bullet N[i] = N[g[i]]\end{aligned}$$

is well defined. ∎

LEMMA 14.9. *Suppose G is a transitive subgroup of $\mathcal{S}_n$. Then for all $N \trianglelefteq G$ the size of an orbit of N must divide n.*

PROOF. Moreover, since the action of G is transitive on $\overline{n}$, it's obvious that the action of G on the orbits of N is also transitive. Therefore each orbit must have the same size and have a total of n elements, hence the result. ∎

THEOREM 14.10. *Suppose that G is a subgroup of $\mathcal{S}_p$ where p is prime which can be generated by elements of order p (i.e. $\underline{p}$–cycles). Then G is simple.*

PROOF. For starters, any such subgroup is transitive since a $\underline{p}$-cycle alone acts transitively on $\overline{p}$.

Let $N \trianglelefteq G$ be a non-trivial normal subgroup of G. We show that $N = G$. Then N must also act transitively since G since the only divisors of p are 1 and p itself (apply the lemma). If the size of N's orbits are all 1, then N must be the trivial group which we assumed it isn't.

By the orbit stabilizer theorem, we also know that p^2 does not divide the order of G (since the stabilizer is a subgroup of $\mathcal{S}_{p-1}$ so its order is not divisible by p). As such, any $\underline{p}$-cycle generates a Sylow p-subgroup of G. By **Sylow2**, all p-groups are conjugate. Therefore any two elements of order p-must be conjugate to a power of any other element of order p. As such N must contain all of the p-cycles in G and since these are assumed to generate G, we must have that $G = N$. ■

The following theorem will be generalized in Theorem 14.13 at the end of the section.

THEOREM 14.11. *$\mathcal{A}_5$ is a simple group.*

PROOF. $\mathcal{A}_5$ is generated by the $\underline{5}$–cycles $(1\,2\,3\,4\,5)$ and $(2\,1\,3\,4\,5)$ (By Exercise 10.75 or Theorem 10.136 back on page 590 and the fact that 5 is a prime number). ■

EXAMPLE 14.12. Define M_{11} to be the group

$$\langle (1\,2\,3\,4\,5\,6\,7\,8\,9\,10\,11), (1\,7\,5\,9\,4\,2\,10\,8\,3\,6\,11) \rangle .$$

Cheating, and using `MAPLE`, we see that this group has order

$$7920 = 11 \cdot 10 \cdot 9 \cdot 8 = 11 \cdot 6!$$

In particular it's order is strictly less than that of $\mathcal{A}_{11}$, and as such, it cannot be $\mathcal{A}_{11}$. In fact, it's not isomorphic to any $\mathcal{A}_n$ since the presence of the 11 as a divisor of the order of M_{11} implies that $n \geq 11$, since no $\mathcal{A}_n$ with $n < 11$ contains an element of order 11.[I] It is called *Mathieu group* 11 after 19th century mathematician Émile Léonard Mathieu.

The above example is the first of what are known as the *sporadic simple groups*. These are simple groups that do not fall into one of the established patterns. There are only 26 such groups. The other simple groups fall into 18 infinite families. The largest sporadic simple group is known as the monster. Its order might give you a hint to why its named this, the order is a whopping

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

which is only slightly larger than an 8 with 53 zeros after it.

EXERCISE 14.1. Show that the monster does not contain an element of order 37 but does contain an element of order 41.

SOLUTION

We now generalize Theorem 14.11 to

THEOREM 14.13. *Suppose $n \neq 4$, then $\mathcal{A}_n$ is simple.*

[I] One way of showing this (if you had several months with nothing better to do) is by showing that M_{11} is sharply four transitive (i.e. acts freely and transitively on distinct quadruples, i.e. elements of $\overline{11}^{*4}$.) That is for any two quadruples of distinct points (a_1, a_2, a_3, a_4) and (b_1, b_2, b_3, b_4) there exists a unique $\sigma \in M_{11}$ so that

$$(\sigma[a_1], \sigma[a_2], \sigma[a_3], \sigma[a_4]) = (b_1, b_2, b_3, b_4).$$

) This would imply that M_{11} and $\overline{11}^{*4}$ have the same number of elements. The latter has

$$\frac{11!}{(11-4)!} = 11 \cdot 10 \cdot 9 \cdot 8$$

elements in it, and so, so does the former.

Proof. {Supply} ∎

Theorem 14.14. *Let p be an odd prime. Suppose that $G \subseteq \mathcal{S}_{p+1}$ is generated by $\underline{p}$–cycles. Then, if G is not simple, then $p+1 = 2^r$ for some positive integer r.*

Proof. Since the trivial group is obviously simple, we assume that G has more than one element. We can assume that G is a transitive subgroup of $\mathcal{S}_n$ for otherwise, G would act transitively on p points (since it contains a p–cycle) and is generated by $\underline{p}$–cycles by assumption, and so is simple by previous theorem.

Now let N be a PNT normal subgroup of G. Since the order of a Sylow p–subgroup of $\mathcal{S}_{p+1}$ has order p, the same is true for G. Thus if we can show that N contains an element of order p (such an element must be a $\underline{p}$–cycle!), just like the previous proof, we can conclude that $N = G$, and we would be done. Thus we make the following assumption:

$$N \text{ does not contain an element of order } p. \tag{1}$$

Now consider $G_{p+1} \preccurlyeq G$, which as always denotes the stabilizer of the letter $p+1$. Note that since G acts transitively on $\overline{p+1}$, and contains a $\underline{p}$–cycle, G_{p+1} acts transitively on $\overline{p}$. Simply conjugate a $\underline{p}$–cycle in G via an element of G which takes the fixed point set of this $\underline{p}$–cycle to $p+1$. Such an element exists by the transitive nature of the action of G. Thus G_{p+1} contains a $\underline{p}$–cycle, and so its action is transitive. (This proves that the action G is 2 transitive.)

Now just like the proof of the previous theorem, we know that since $N \cap G_{p+1} \trianglelefteq G_{p+1}$, this intersection must either be trivial or act transitively on $\overline{p}$ as well. But if the action is transitive, then by OST we know that p divides the order of this intersection. But by Cauchy's lemma, this means that this intersection would contain an element of order p. Hence N would contain an element of order p, contradicting (1). Therefore $N \cap G_{p+1}$ is trivial. Moreover, either by using the fact that N is normal, or making a similar argument we can see for all $i \in \overline{p+1}$, $N \cap G_i$ is also trivial.

Next, note that N must also act transitively on $\overline{p+1}$. This is because if not, we could divide $\overline{p+1}$ up into more than 1 orbit of N's action on $\overline{p+1}$,

$$\overline{p+1}/_N = \{\mathbf{O}_1, \ldots, \mathbf{O}_m\}$$

with $m > 1$. And, since $N \trianglelefteq G$, G would admit a well defined action on these orbits. Note that since N only trivially meets the stabilizer of every point, each of these orbits must have more than 1 letter. Now suppose that $a_1, a_2 \in \mathbf{O}_1$ and $b \in \mathbf{O}_2$. Then since the action of G on $\overline{p+1}$ is 2–transitive, there exists a $g \in G$ so that $g[a_1] = a_1$ and $g[a_2] = b$. This contradicts the fact that G's action on $\overline{p+1}/_N$ is well defined. Thus N must act transitively on $\overline{p+1}$.

Thus we have proved that $N \circlearrowright \overline{p+1}$ is free and transitive. In particular, by OST,

$$\frac{|N|}{|N_{p+1}|} = \#(N[(p+1)]) \Rightarrow |N| = p+1.^{\text{II}}$$

($|N_{p+1}| = 1$, $\#(N[(p+1)]) = p+1$)

Moving on, let's denote a $\underline{p}$–cycle in G by $\mathfrak{p}$. As already pointed out, we can assume that $\mathfrak{p}[(p+1)] = p+1$, i.e that $\mathfrak{p} \in G_{p+1}$. We know that $\mathfrak{p} \in \mathcal{S}_{p+1}$ commutes with only powers of itself. Thus the same must be true in the smaller group G. In particular, since N has already been proven to not meet G_{p+1}, $\mathfrak{p}$ cannot commute with any element of N. Moreover the same can be said for any power of $\mathfrak{p}$. This means for any two non-identity elements of N, n_1, n_2, there exists an element of $\mathfrak{p}^i \in \langle \mathfrak{p} \rangle$ so that n_1 is conjugate to n_2 via $\mathfrak{p}^i$. In particular, all elements of N must be conjugate (in G), and since conjugate elements have the same order, every non-identity element of N must have the same order. This is only possible if that common order is a prime q, since every (finite) group must have an element of prime order. In particular N must be a q–group, i.e. the order of N must be p^r for some integer r. But since

$$p+1 = |N| = q^r$$

we must have that q is even, for otherwise q^r would be odd while $p+1$ would be even. But the only even prime number is 2. Hence $q = 2$ and $p+1 = q^r$ as claimed. ∎

Note we have actually proven something a bit stronger:

Corollary 14.15. *With the same hypothesis as in the theorem, any normal subgroup of G must be isomorphic to $\mathbb{Z}_2^r$ where r is taken so that $2^r = |G|$. (So r is log base 2 of $|G|$.)*

[II] In general, if a group acts freely and transitively on a set, then the group and the set must have the same order.

Proof. In the proof, we ended up showing that every element of N has order 2. But any group with this property must be isomorphic to $\mathbb{Z}_2^r$ for some integer. The specifics of the integer r follow from the fact that we also proved that the order of N must be $p+1$. ■

Example 14.16. This makes it even easier to come up with examples of simple groups. Simply define a subgroup of say, $\mathcal{S}_{12}$, generated by 11–cycles. Of course this isn't as cool as we would hope since most of the time the group we end up defining is simply $\mathcal{S}_{11}$. In fact, it's not hard to see that the only simple subgroup of $\mathcal{S}_6$ is $\mathcal{A}_6$ and $\mathcal{A}_5$. Thus we are not really doing anything very exciting in most cases.

One case where things get a bit interesting is for the groups $\prod p$ considered in the quotient groups chapter. These groups all act on a set with $p+1$ points (namely the lines in $\mathbb{Z}_p^2$) and are generated by the two elements, $\mathfrak{L}$ and $\mathfrak{R}$, which have order p. Thus we have proven that these groups are simple for say $p = 5, 11, 13, 17, 19$, etc.

However for other primes, where the theorem doesn't directly apply, we can use the corollary. This is because for p a prime of the form $2^r - 1$ (for $r > 0$) we would have that 2^r is the order of a Sylow 2–subgroup of $\prod p$. This is because $\prod p$ has order $(p+1)p(p-1)/2$ for odd primes, and if $p + 1 = 2^r$, then $p - 1$ can only be divisible by 2 once (four cannot divide two integers whose difference is 2). Thus the 2 in the denominator kills this sole 2 in $p - 1$, and we are left only with the r $2's$ in $p + 1$. However, with some effort, we could see that a Sylow 2–subgroup of $\prod p$, with $p = 2^r - 1$, is isomorphic to $\mathcal{D}_{2^{(r-1)}}$.[III] The only time that $\mathcal{D}_{2^{(r-1)}}$ is abelian (and can be isomorphic to $\mathbb{Z}_2^r$) is when $r = 1$ or when $r = 2$, i.e. when $p = 1$ or $p = 3$.

We summarize this example:

Example /Theorem 14.17. For p a prime not equal to 2 or 3, $\prod p$ is a simple group. ■

In fact, it is also the case that $\mathbf{PSL}_n\left(\mathbb{Z}_p\right)$ is simple in all cases except for the cases that $n = 2$ and $p = 2$ or $p = 3$.

14.1 Simple Groups of Order Less than 100

(Non–abelian) simple groups are rather sparse in group theory. Of all groups of order less than 168, only one, $\mathcal{A}_6$, is simple. The next smallest (non–abelian) simple group is $\mathbf{GL}_2\left(\mathbb{Z}_2\right)$, which happens to be isomorphic to $\mathbf{PSL}_2\left(\mathbb{Z}_7\right)$, and has order 168.

Checking out the array, we see that we've already limited many possibilities from general theory. Here, numbers colored in *red* are powers of a prime, so every group of that order is a p–group (and can't be non–abelian and simple since such groups have a non-trivial center). Groups of an order colored in *blue* are the product of either two or three prime numbers. We took care of groups whose order is the product of two primes in the last chapter where we showed that such groups are either cyclic, or a semi–direct product. With a little maneuvering, proving that no group whose order is the product of three distinct primes is also not too bad.

Exercise 14.2. Suppose that G is a group of order pqr, where $p < q < r$ are primes. Prove that a Sylow r–subgroup of G is normal.

Hint! First prove that if not, then by a counting argument that both the Sylow p and q–subgroups of G must be normal. Prove that $\left\langle P_p, P_q \right\rangle$ is isomorphic to $\mathbb{Z}_p \times \mathbb{Z}_q$. Prove that a Sylow r–subgroup P_r must be normal by studying the way that it can act on $\mathbb{Z}_p \times \mathbb{Z}_q$ by automorphisms.

SOLUTION

Groups colored in green have the form $p^r q^s$ in such a way that none of the powers of q less than s are equivalent to 1 mod p or none of the powers of p less than r are equivalent to 1 mod q. Numbers in violet are of the form $2^r \cdot 3$ for $r > 1$.

Theorem 14.18. *There are no simple groups of order* $2^r \cdot 3$ $(r > 0)$.

[III] We only need to prove that the Sylow 2–subgroup of $\prod p$ is not isomorphic to $\mathbb{Z}_2^r$ in these cases, and there are several ways of doing this. One is to show that it's not abelian by finding elements of order 2 in $\prod p$ which normalize each other but don't commute. Another is to show that there exists an element of order 2^r in $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$, which therefore descend to elements of order 2^{r-1} in $\prod p$!

Proof. Let G be a group of order $n := 2^r \cdot 3$. If $r = 3$, then $n = 6$ and we know that such a group cannot be simple (either by noting that this means that $G \cong \mathcal{S}_2$ or $\mathbb{Z}_6$, or simply by using the fact that 6 is the product of just two primes.) If $r > 1$, than $n > 6$. Now if P_2, a Sylow 2–subgroup of G, is not normal in G, then $n_2(G) = 3$. Therefore, since the action of $G \circlearrowright \mathbf{Syl}_2 G$ is transitive (and hence non-trivial) we have a non-trivial homomorphism $G \to \mathcal{S}_2$. But since $|G| > 6$ this homomorphism cannot be injective, and so the kernel of this homomorphism gives us a PNT, normal subgroup of G, proving that G is not simple. ∎

I colored the number 60 in yellow since we know already that there is a simple group of order 60, namely $\mathcal{A}_5$, however we still need to check that this is the only simple group of order 60.

	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	

This leaves only a handful of numbers that we need to evaluate. I imply for every part of the below, when dealing with the integer

$$n = p_1^{n_1} \cdots p_m^{n_m}.$$

Let G be a group of order n and for each $i \in \overline{m}$ let P_{p_i} be a Sylow p_i–subgroup of G.

For instance, the item labeled **56** below is implied to start with the statement: let G be a group of order 56 with P_7 a Sylow 7–subgroup of G, and P_2 a Sylow 2–subgroup of G.

We use the fact that the action of a group G on its Sylow p–subgroups is transitive, and thus yields a non-trivial homomorphism

$$G \to \mathcal{S}_{n_p(G)}.$$

(1) (36)

Exercise 14.3. Prove that no group of order 36 is simple.

Solution

(2) (56) The number 56 factors as $7 \cdot 2^3$. If P_7 is not normal in G, then, as the only divisor, besides 1, of 56 which is equivalent to 1 mod 7, $n_7(G) = 8$ by **Sylow3**. This means there are $6 \cdot 8 = 48$ elements of order 7 in G. This leaves only 8 elements of G which do not have order 7. This is only enough to from a single Sylow 2–subgroup of G, i.e. $n_2(G) = 1$ and $P_2 \trianglelefteq G$ gives us a PNT normal subgroup of G, proving that G is not simple.

(3) (60) The number 60 factors as $2^2 \cdot 3 \cdot 5$. This number is quite a bit different from the others. We need to prove that if G is a simple group of order 60, then $G \cong \mathcal{A}_5$. So we assume that G is a simple group of order 60 which is not isomorphic to $\mathcal{A}_5$. Like a lying child, we string together a bunch of lies (i.e. false assumptions) which collapse like a deck of our cards leading us all the way back to prove that the assumption that G is not isomorphic to $\mathcal{A}_5$ is false.

Using the assumption that G is simple and laying with **Sylow3** for a bit, we see we have $n_2(G) = 5$ or 15, $n_3(G) = 4$ or 10, and $n_5(G) = 6$.

If $n_3(G) = 4$ we are done (why?).

If $n_2(G) = 5$, we are also done. This is because we would have a non-trivial, injective (since G is assumed simple) homomorphism $G \to \mathcal{S}_5$. The image of this homomorphism must be exactly $\mathcal{A}_5$ since otherwise, the intersection of this image with $\mathcal{A}_5$ would give us a PNT normal subgroup of the simple group G, which is impossible the very definition of simple. Therefore, in that case, $G \cong \mathcal{A}_5$.

There is only one possibility left: the case that $n_2(G) = 15$, $n_3(G) = 10$, and $n_5(G) = 6$.

As a group of order 4, there are two possibilities for P_2: P_2 could be isomorphic to $\mathbb{Z}_4$ or to $\mathbb{Z}_2^2$. In the first case, we would have $15 \cdot 2 = 30$ elements of order 4 in G, $10 \cdot 2 = 20$ elements of order 3 in G, and $6 \cdot 4 = 24$ elements of order 5 in G. But $30 + 20 + 24 = 84 > 60$, and (as Bob Barker would say) we are over! Therefore, P_2 cannot be isomorphic to $\mathbb{Z}_4$ and so we are left with the case that $P_4 \cong \mathbb{Z}_2^2$.

Let's summarize, we only need to deal with the case that $n_2(G) = 15$ and $P_2 \cong \mathbb{Z}_2^2$.

I claim that if P_2' is another Sylow 2–subgroup of G, then $P_2 \cap P_2'$ is trivial. This would mean $15 \cdot 3 = 45$ elements of order 2 in G pushing the total number of elements of G up to $45 + 20 + 24 = 89$, and we are now (even more) over our limit, which is 60. So the end is in sight: we need to prove that $P_2 \cap P_2'$ is trivial.

We set up our final straw man and assume that $P := P_2 \cap P_2'$ is non-trivial. Let's denote

$$N := \mathrm{N}_G(P),$$

the normalizer of this intersection in G. Since G is assumed simple, and P assumed non-trivial, $N \neq G$, for otherwise P would be a PNT normal subgroup of G (a subgroup of a group is normal iff its normalizer is the entire group).

Since P_2 and P_2' are abelian, we know that P is normal in both. Therefore, both are in N, and so

$$P_2P_2' = \{g_1g_2 : g_1 \in P_2, g_2 \in P_2'\}$$

is also contained in N. Proposition 10.102 says that P_2P_2' contains

$$\frac{|P_1||P_2|}{\cdot}|P| = \frac{4 \cdot 4}{2} = 8$$

elements.[1] Therefore $|N| \Big| |G|$ (by Lagrange) and $8 < |N| < 60$ (it cannot be exactly 8 since 8 does not divide 60).

This only leaves us with two possibilities: the first is that $|N| = 12$ and the second is that $|N| = 20$.

If N has order 12, then N would have index $60/12 = 5$ and the action of G on the 5 cosets, G/N, would have a permutation representation which would give us an isomorphism of G with $\mathcal{A}_5$. (Fill in the details!)

On the other hand, if $|N| = 20$, we would have an index $60/20 = 3$ subgroup of G, and the coset action of G on the set of cosets, G/N, would give us a non-trivial homomorphism $G \to \mathcal{S}_3$ which is impossible since G is assumed simple and $\mathcal{S}_3$ has only 6 elements.

This means that we have contradicted the fact that $P_2 \cap P_2'$ is non-trivial, which contradicts the assumption that $n_2(G) = 15$, which contradicts the assumption that G is not isomorphic to $\mathcal{A}_5$.

(4) (72) The number 72 factors as $72 = 3^2 \cdot 2^3$. Like in the case that $n = 48$, if P_3 is not normal in G, then $n_3(G) = 4$. This gives rise to a non-trivial homomorphism $G \to \mathcal{S}_4$ which cannot be injective since $72 > 24$.

(5) (80) The number 80 factors as $2^4 \cdot 5$. Therefore, if P_5 is not normal in G, then $n_5(G) = 16$ (this is the only divisor of 80, besides 1, which is equivalent to 1 mod 5). Like in the case that $n = 56$, this would mean that $P_2 \trianglelefteq G$.

(6) (84) The number 84 factors as $2^2 \cdot 3 \cdot 7$. Now the divisors of $2^2 \cdot 3$ are 1, 2, 4, 3, and 12 which equal 1, 2, 4, 3, and 5 mod 7. In particular, none are equivalent to 1 mod 7 and so $n_7(G) = 1$, and so P_7 is a non-trivial normal subgroup of G and G is not simple.

[1] Why doesn't the fact that 8 divide 60 contradict Lagrange's theorem?

14.2 Quick Facts

To get some intuition about how common simple groups are we give you this table listing the orders and descriptions of the smallest non-ableian simple groups.

Order	60	168	360	504	660	1092	2448	2520
=	$5 \cdot 3 \cdot 2^2$	$7 \cdot 3 \cdot 2^3$	$5 \cdot 3^2 \cdot 2^3$	$7 \cdot 3^2 \cdot 2^3$	$11 \cdot 5 \cdot 3 \cdot 2^2$	$13 \cdot 7 \cdot 3 \cdot 2^2$	$17 \cdot 3^2 \cdot 2^4$	$7 \cdot 5 \cdot 3^2 \cdot 2^3$
Group	$\mathcal{A}_5$	$\mathbf{SL}_3(\mathbb{Z}_2)$	$\mathcal{A}_6$	$\mathbf{SL}_2(\mathbb{F}_8)$	$\mathbf{PSL}_2(\mathbb{Z}_{11})$	$\mathbf{PSL}_2(\mathbb{Z}_{13})$	$\mathbf{PSL}_2(\mathbb{Z}_{17})$	$\mathcal{A}_7$

We could also describe $\mathcal{A}_5$ as $\mathbf{PSL}_2(\mathbb{Z}_5)$ or as $S(\text{✶}^3)$, the symmetries of the dodecahedron (or icosahedron). And we could have described $\mathbf{SL}_3(\mathbb{Z}_2)$ as $\mathbf{PSL}_2(\mathbb{Z}_7)$ (this isomorphism is non-trivial, can you prove it?). The one set that we haven't encountered before is the set $\mathbb{F}_8$. It is a field like $\mathbb{Z}_p$ in the terms of the fact that it's a place where one can add, subtract, multiply, and divide (except by zero). However, it has 8 elements instead of a prime number of elements (it certainly isn't $\mathbb{Z}_8$ since you can't divide by 2,4, or 6 there). Check out Example E.67 in the appendix for a further discussion of this along with an addition and multiplication table.

14.3 More on Transitive Groups

Lemma 14.19. *Suppose that $G \leqslant \mathcal{S}_5$ has order > 24. Then G is either $\mathcal{A}_5$ or all of $\mathcal{S}_5$.*

Proof. Using simplicity we've seen that proper subgroup of $\mathcal{A}_5$ can have order larger than 12. Now for G a subgroup $\mathcal{S}_5$ which is not contained in $\mathcal{A}_5$ we have that

$$|\mathcal{S}_5| = \frac{|G| \cdot |\mathcal{A}_5|}{|G \cap \mathcal{A}_5|} \Rightarrow |G| = 2 \cdot |G \cap \mathcal{A}_5| \leq 2 \cdot 12 = 24.$$

■

Remark 14.20. Since $\mathcal{S}_4 \leqslant \mathcal{S}_5$ (which has order 24), this inequality is exact.

Lemma 14.21. *Suppose that $G \leqslant \mathcal{S}_5$ contains a $\underline{5}$–cycle which does not generate a normal subgroup of G. Then $G = \mathcal{A}_5$ or $G = \mathcal{S}_4$.*

Proof. Assume that the group generated by a $\underline{5}$–cycle in G is not normal in G. Since 25 does not divide $|\mathcal{S}_5| = 5!$, we know that 25 also does not divide $|G|$. Hence that $\underline{5}$–cycle in G generates a Sylow 5–subgroup of G. Since it does not generate a normal one, then there must be at least 6 of them by Sylow 3. Therefore there is at least $6 \cdot 4 = 24$ elements of order 5 in G. Since 5 does not divide 24 there has to be even more elements than that in G. But the previous lemma showed that this implies that $G = \mathcal{A}_5$ or $\mathcal{S}_5$ as desired. ■

In the theory of transitive subgroups of $\mathcal{S}_n$, clearly it would be silly to study conjugate subgroups as if they were really the same thing. This is because in $\mathcal{S}_n$ conjugation is replacement, i.e. renaming. Therefore two conjugate subgroups of $\mathcal{S}_n$ simply differ by the ordering of the set $\overline{n} = \{1, 2, \ldots, n\}$.

Theorem 14.22. *Every transitive subgroup of $\mathcal{S}_5$ is conjugate (hence isomorphic to) one of the following*

$$\mathcal{S}_5 \quad \mathcal{A}_5 \quad \langle (1\,2\,3\,4\,5), (1\,2\,4\,3) \rangle \quad \langle (1\,2\,3\,4\,5), (1\,4)(2\,3) \rangle \quad \langle (1\,2\,3\,4\,5) \rangle.$$

Remark 14.23. Note that the 3rd group listed could be described as $\mathrm{N}_{\mathcal{S}_5}\langle (1\,2\,3\,4\,5) \rangle$, the 4th as $\mathrm{N}_{\mathcal{A}_5}\langle (1\,2\,3\,4\,5) \rangle$, which is isomorphic to $\mathcal{D}_5$, and the 5th is (obviously) isomorphic to $\mathbb{Z}_5$.

Proof. Let G be a transitive subgroup of $\mathcal{S}_5$ which we assume is not $\mathcal{A}_5$ or $\mathcal{S}_5$. Since G acts transitively on 5 things, OST applies to show that $5 \Big| |G|$. As the only elements of $\mathcal{S}_5$ of order 5, G must contain a $\underline{5}$–cycle, σ, which after conjugating G we can assume $\sigma = (1\,2\,3\,4\,5)$. By the previous lemma we know that $\langle \sigma \rangle$ is a normal subgroup of G. Hence G must be contained inside of the normalizer

$$N := \langle (1\,2\,3\,4\,5), (1\,2\,4\,3) \rangle$$

of $(1\,2\,3\,4\,5)$. Note that this normalizer has order 20. Since $5 \Big| |G|$ and $|G| \Big| 20$, we are left with the following possibilities:

(1) $|G| = 20$ in which case $G = N$.

(2) $|G| = 5$, in which case $G = \langle (1\,2\,3\,4\,5) \rangle$.

(3) $|G| = 10$, in which case it is plain to see that $G = \langle (1\,2\,3\,4\,5), (1\,4)(2\,3) \rangle$. Indeed there are several ways of doing this. I'll give you one of the less fancy. For starters now that our assumption is that G has index 2 in N. Therefore, taking the quotient N/G (which makes sense since $|G|$ has index 2) we get a homomorphism from $N \to \mathbb{Z}_2$ with kernel G. But since 5, the order of $(1\,2\,3\,4\,5)$ and 2 are relatively prime, the $\underline{5}$–cycle must be in the kernel, i.e. in G (well, we knew this already), so must $(1\,2\,4\,3)$ go to the non-trivial element of $\mathbb{Z}_2$, so it's square must also be in the kernel. But it's square is $(1\,4)(2\,3)$. Therefore $\langle (1\,2\,3\,4\,5), (1\,4)(2\,3) \rangle \leqslant G$ and equality must hold since each have order 10.

■

Corollary 14.24. *The only simple groups which contain an index 5 subgroup are $\mathbb{Z}_5$ and $\mathcal{A}_5$.*

Proof. Suppose G is a simple group with an index 5 subgroup H. Then the coset action $G \circlearrowright G/H$ has a permutation representation $G \to \mathcal{S}_5$ onto a transitive subgroup of $\mathcal{S}_5$. Since G is simple this representation must be injective. Therefore G must be isomorphic to one of the five groups of the theorem. Clearly the only simple groups of these 5 are $\mathcal{A}_5$ and $\mathbb{Z}_5$. ■

Note that almost nothing we've discussed here works when we attempt to examine transitive subgroups of $\mathcal{S}_6$ up to conjugacy. For instance we don't even know that such subgroups contain a $\underline{6}$–cycle, it needn't even have an element of order 6 at all! For instance if consider the malnormal subgroup of order 4 and index 6, $H := \langle (1\,2\,3\,4) \rangle$ of $\mathcal{S}_4$. Then the coset action $\mathcal{S}_4 \circlearrowright \mathcal{S}_4/H$ is transitive (like all coset actions are) and so after choosing a dictionary on these 6 cosets we see there is a transitive subgroup of $\mathcal{S}_6$ which is isomorphic to $\mathcal{S}_4$. But $\mathcal{S}_4$ does not contain an element of order 6, so we have found a transitive subgroup of $\mathcal{S}_6$ which does not contain an element of order 6 or even 5.

The same trick even works if instead we consider $\mathcal{A}_4$ and H the malnormal subgroup of index 6 which is generated by any of the $\underline{2}\,\underline{2}$–cycles in $\mathcal{A}_4$. Since $\mathcal{A}_4$ does not even contain an element of order 4, this gives us a transitive subgroup of $\mathcal{S}_6$ without an element of order 4, 5, or 6!

Remark 14.25. The first example of $\mathcal{S}_4 \leqslant \mathcal{S}_6$ can be given by sending generators

$$(1\,2\,3\,4) \mapsto (2\,3\,4\,5) \quad \text{and} \quad (1\,2\,3) \mapsto (1\,2\,3)(4\,5\,6).$$

The second example of $\mathcal{A}_4 \leqslant \mathcal{S}_6$ can be given by sending generators

$$(1\,3)(2\,4) \mapsto (2\,4)(3\,5) \quad \text{and} \quad (1\,2\,3) \mapsto (1\,2\,3)(4\,5\,6).$$

Actually both can be realized geometrically by identifying $\mathcal{S}_4$ with the symmetries of a cube, $S(\blacklozenge^3)$ and then checking out the action of the symmetries of a cube on the six faces of the cube.

However this is the worst things can get since if $6 \Big| |G|$ implies that both 2 and 3 divide $|G|$ and so by Cauchy's lemma, there must be elements of order 2 and 3 in G.

Exercise 14.4. Represent $\mathcal{A}_4 \times \mathbb{Z}_2$ as a transitive subgroup of $\mathcal{S}_6$.
Hint! Consider the coset action

$$\mathcal{A}_4 \times \mathbb{Z}_2 \circlearrowright {}^{\mathcal{A}_4 \times \mathbb{Z}_2}\!/_{H}$$

where H is

$$H := \langle ((1\,2)(3\,4), [1]_2), ((1\,3)(2\,4), [1]_2) \rangle \preccurlyeq \mathcal{A}_4 \times \mathbb{Z}_2.$$

Exercise 14.5. Represent $\mathcal{D}_6$ as a transitive subgroup of $\mathcal{S}_6$ in two non–conjugate ways.
Hint! Geometrically you can see this via the isomorphism $\mathcal{D}_6 \cong \mathcal{D}_3 \times \mathbb{Z}_2$. The first term, $\mathcal{D}_6$ is geometrically the symmetries of a hexagon, while the second term, $\mathcal{D}_3 \times \mathbb{Z}_2$ can be seen as symmetries of a prism.

It turns out that up to conjugacy there are a total of 16 transitive subgroups of $\mathcal{S}_6$.

14.4 $\mathbf{SO}_3(\mathbb{R})$ IS SIMPLE*

We prove that the (uncountably) infinite group of rotations (in $\mathbb{R}^3$), $\mathbf{SO}_3(\mathbb{R})$, is simple.

We do this by working in the easier (I guess I can't safely say simpler!) group $\mathbb{S}^3$. Note that if we can show that the only normal subgroup of $\mathbb{S}^3$ is $\{-1, +1\}$ we are going to be done. This uses 11.36 and the fact that $\mathbf{SO}_3(\mathbb{R}) \cong \mathbb{S}^3/\{+1, -1\}$.

We need to prove that if $N \trianglelefteq \mathbb{S}^3$ contains a quaternion with non-zero pure part, then $N = \mathbb{S}^3$.

Recall that $\mathbb{S}^3$ is foliated by spheres or radius corresponding to the length of the pure part of the quaternions $a + \mathbf{v}$. If we write this quaternion in the standard form for a unit quaternion

$$\cos\theta + \sin\theta\,\widehat{\mathbf{u}} \in \mathbb{S}^2$$

then the sphere, denoted $\mathbb{S}^2_{\cos\theta}$, has radius $\sin\theta$. Of course $\mathbb{S}^2_1$ contains only 1 and $\mathbb{S}^2_{-1}$ contains only -1. With this notation $\mathbb{S}^2_o$ is what we've been calling $\mathbb{S}^2$, i.e. it's the set of pure unit quaternions.

We've seen that these 2-spheres make up the conjugacy classes {where} of elements in $\mathbb{S}^3$, i.e. two elements are conjugate iff they are in the same sphere. Thus if N contains just one element of a given 2-sphere it must contain all elements of that given 2-sphere.

Recall further that {where} any element of $\mathbb{S}^3$ can be written as the product of two elements of $\mathbb{S}^2 = \mathbb{S}^2_0$. Therefore, if we can prove that N contains a single pure quaternion, we're going to be guaranteed that it contains all unit pure quaternions and (by closure of N under multiplication) all of $\mathbb{S}^3$.

Therefore our goal is to prove that if N contains a quaternion with a non-zero pure part, say $\mathsf{q} = \cos\theta + \sin\theta\widehat{\mathbf{u}}$ with $0 < \theta < \pi$ (i.e. q is not equal to 1 or -1), then N contains a pure unit quaternion. There are two cases:

- If θ is an angle $\pi/4$ and $3\pi/4$, then $\cot\theta$ is between -1 and 1. This means that there is some angle ψ which solves this equation

$$0 = \cos^2\theta - \sin^2\theta \cdot \cos\psi. \tag{2}$$

  This means that if we take $\mathsf{p} = \cos\theta + \sin\theta\,\widehat{\mathbf{u}}'$ (which as an element conjugate to q is also in the normal subgroup, N) as any quaternion which is an angle ψ away from q, then by closure the product qp is also in N and thanks to our choice of ψ,

$$\begin{aligned}
\mathbf{Re}(\mathsf{p} * \mathsf{q}) &= \mathbf{Re}((\cos\theta + \sin\theta\,\widehat{\mathbf{u}})(\cos\theta + \sin\theta\,\widehat{\mathbf{u}}')) \\
&= \cos^2\theta - \sin^2\theta\,\widehat{\mathbf{u}} \bullet \widehat{\mathbf{u}}' && (\mathbb{R} \times \mathbb{R}^3 \text{ quat mult}) \\
&= \cos^2\theta - \sin^2\theta\,\cos\psi && (\theta_{\widehat{\mathbf{u}},\widehat{\mathbf{u}}'} = \psi \text{ and } \left\|\widehat{\mathbf{u}}\right\| = \left\|\widehat{\mathbf{u}}'\right\| = 1) \\
&= 0. && (\text{Equation } (2))
\end{aligned}$$

  This means that this product, qp has no real part, i.e. this product is pure. Therefore, in this product, we have found a pure element of N and we are done.

- Now if the angle θ doesn't fall in the specified range, the above trick doesn't work (indeed, the argument above also shows that there is no element in the same sphere of q whose product is pure). However, we can clearly find a multiple of θ which is between the specified range. Indeed, if θ is in the first bad range, i.e. it's between 0 and $\pi/4$, take the smallest integer n so that $n\theta \geq \pi/4$ (this exists and works due to the division algorith). If its in the second bad range, i.e. if it's between $3\pi/4$ and π, then a bit of fudging allows the same basic argument to work.[1] Then $\mathsf{q}^n = \cos n\theta + \sin n\theta\widehat{\mathbf{u}}$ and $\mathsf{q}^n \in N$ by closure. Lastly, the choice of n shows us that that this power falls into the first case.

We have proven:

THEOREM 14.26. *The group of 3 dimensional rotations* (i.e. the group $\mathbf{SO}_3(\mathbb{R})$) is simple. ■

[1] For example, if $\theta = 17\pi/5014$, then take n as 74.

15 More Odds and Ends

An automorphism of a poset $(X, \preccurlyeq)$ is a bijection $\varphi\colon X \to X$ which preserves order. That is

$$\varphi(x) \preccurlyeq \varphi(y)$$

whenever $x \preccurlyeq y$. The set of all such automorphisms is denoted Aut(X, $\preccurlyeq$).

Exercise 15.1.

(1) Show that Aut $(X, \preccurlyeq)$ is a group.

(2) Let n be a positive integer and let $\widehat{n}$ be the poset of divisors of n. Prove as much as you can about the automorphisms of this poset. Work some specific examples, say for the integers 13, 860, 256, and 6. For instance give a condition for these groups to be isomorphic. What are the global fixed point sets look like? For instance consider the example $n = 1420$.

(3) Define a map ι which takes a divisor $d \in \widehat{n}$ to n/d. Show that ι is not an automorphism (except in trivial situations) but satisfies a different, but similar condition. Define, in a reasonable way, a new group which contains all automorphisms and ι. What is the index of the set of all automorphisms in this group? Show that the set of all automorphisms is normal in this new group. What condition on n does one need for there to exist a fixed point for the bijection ι?

SOLUTION

The following theorem generalizes Theorem 10.61 back on page 529 which says that an index two subgroup of a group is normal in that group.

Theorem 15.1. *Suppose that $H \preccurlyeq G$ has index the smallest (positive) prime divisor of G. Then $H \trianglelefteq G$.*

Proof. Let's denote the index of H in G by p. Consider the coset action, $G \circlearrowright G/H$. Now the restricted action, $H \circlearrowright G/H$, has a global fixed point, namely the trivial coset H. This is because every $h \in H$ we have $h \bullet H = H$. Therefore H acts on the set of all other cosets, $G/H \setminus H := X$. Since $\#X = p - 1$, and p is assumed to be less than every other prime divisor of $|G|$, we know that every prime divisor of $|\mathcal{S}_X| = (p! - 1)$ is strictly less than every prime divisor of $|G|$. In particular, $|H|$ and $|\mathcal{S}_X|$ are relatively prime. Therefore, the permutation representation

$$H \to \mathcal{S}_X$$

(like every homomorphism of groups of relatively prime order) has trivial image. This proves that $H \trianglelefteq G$, c.f. Remark 10.65.**??** on page **??**. ∎

The reason why this generalizes Theorem 10.61 back on page 529 is because 2, as the smallest prime, is also the smallest prime of any number which it divides.

Corollary 15.2. *Any subgroup of a p–group P of index p is normal in P.*

Proof. Since p is the only prime divisor of $|P|$, it must be the smallest prime divisor of $|P|$. Therefore the theorem applies. ∎

EXERCISE 15.2. Find the order of the group generated by the following 4 permutations:

$$\sigma_1 := (1\,5\,9) \qquad \sigma_2 := (1\,4\,7\,10)(2\,5\,8\,11)(3\,6\,9\,12)$$
$$\sigma_3 := (1\,5)(2\,10)(4\,8)(7\,11) \qquad \sigma_4 := (1\,7)(3\,9)(5\,11).$$

SOLUTION

Hint! (As usual, G denotes the entire group). Define $G_0 := G$, $G_1 := \langle \sigma_1, \sigma_2, \sigma_3 \rangle$, and $G_2 := \langle \sigma_1, \sigma_2 \rangle$. Prove and use that σ_4 normalizes G_1 but $\sigma_4 \notin G_1$, and σ_3 normalizes G_2 but $\sigma_3 \notin G_2$. Then prove and use that σ_2 normalizes

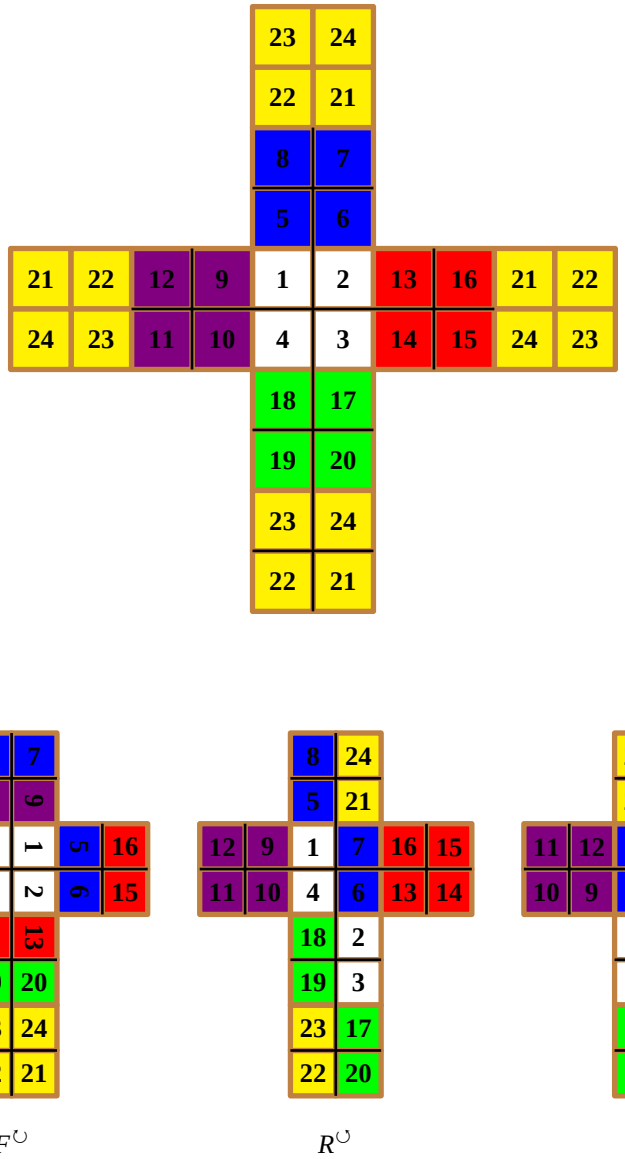
$$\left\langle \sigma_1,\ \sigma_2\sigma_1\sigma_2^{-1},\ \sigma_2^2\sigma_1\sigma_2^{-2},\ \sigma_2^3\sigma_1\sigma_2^{-3} \right\rangle$$

but is not in there.

15.1 More On Rubik's Cube

Let's study the Rubik's cube group, $\mathcal{RK}_2$, first presented on page 439. It is the group of positions (really of motions) of the $2 \times 2 \times 2$ Rubik's cube. In this section we prove, by hand, that there are $88,179,840$ elements of this group. Said in another way, if you were to pick up an unsolved Rubik's cube and randomly do something to it, you would have a 1 in $88,179,840$ chance of solving it.

In order to see it better, let's flatten it out and label the faces (which allows us to think of $\mathcal{RK}_2 \preccurlyeq \mathcal{S}_{24}$ explicitly).



$F^{\circlearrowleft}$ $R^{\circlearrowleft}$ $L^{\circlearrowleft}$

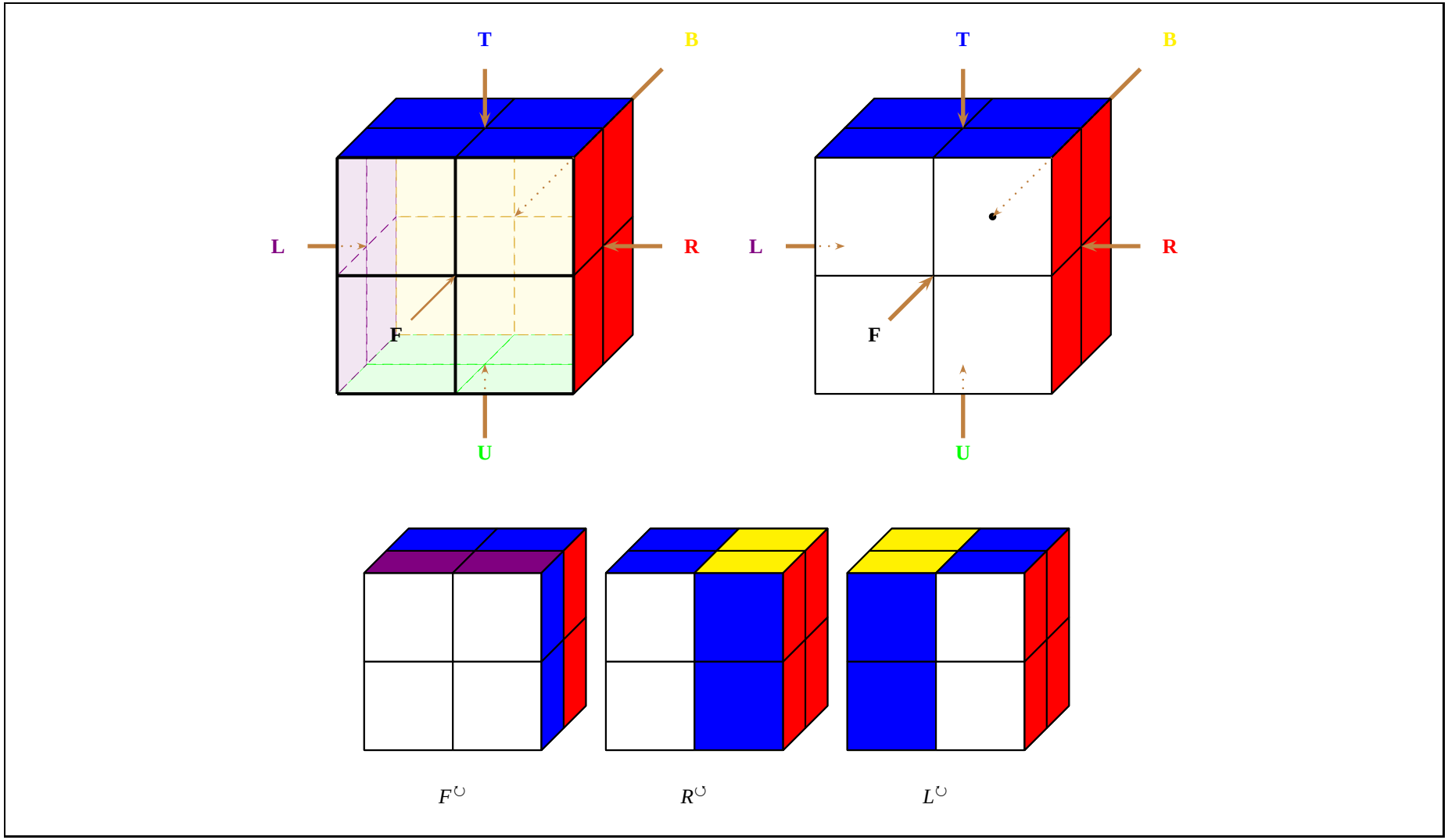


Figure 15.1: A $2 \times 2 \times 2$ Rubik's cube along with (the end result of applying) three basic moves. The faces are labeled F=Front, L=Left, R=Right, B=Back (not bottom!), U=Under, T=Top.

This means, as permutations, the three depicted permutations are

$$F^{\circlearrowleft} = (1\ 2\ 3\ 4)(5\ 13\ 17\ 10)(6\ 14\ 18\ 9)$$
$$R^{\circlearrowleft} = (2\ 17\ 24\ 7)(3\ 20\ 21\ 6)(13\ 14\ 15\ 16)$$
$$L^{\circlearrowleft} = (1\ 18\ 23\ 8)(4\ 19\ 22\ 5)(9\ 10\ 11\ 12)$$

or color coding them:

$$F^{\circlearrowleft} = (1\ 2\ 3\ 4)(5\ 13\ 17\ 10)(6\ 14\ 18\ 9)$$
$$R^{\circlearrowleft} = (2\ 17\ 24\ 7)(3\ 20\ 21\ 6)(13\ 14\ 15\ 16)$$
$$L^{\circlearrowleft} = (1\ 18\ 23\ 8)(4\ 19\ 22\ 5)(9\ 10\ 11\ 12)$$

I claim that these three moves generate all of $\mathcal{RK}_2$. Obviously

$$F^{\circlearrowleft} \quad R^{\circlearrowleft} \quad L^{\circlearrowleft} \quad T^{\circlearrowleft} \quad B^{\circlearrowleft} \quad U^{\circlearrowleft}$$

generate $\mathcal{RK}_2$. Indeed this is what you do when you solve a Rubik's cube, you just move one face at a time. (Of course we can rotate the face backwards by rotating it foward 3 times so the fact that we listed only turns in one direction shouldn't scare you!) However we can get the additional three moves, $U^{\circlearrowleft}$, $T^{\circlearrowleft}$, and $B^{\circlearrowleft}$ by using $R^{\circlearrowleft}$ and $L^{\circlearrowleft}$ together to move the relevant face to the front and then performing $F^{\circlearrowleft}$ and then moving the face back, in other words by conjugating. Therefore:

$$T^{\circlearrowleft} = (R^{\circlearrowleft} \circ L^{\circlearrowleft})^{-1} \circ F^{\circlearrowleft} \circ (R^{\circlearrowleft} \circ L^{\circlearrowleft})$$
$$B^{\circlearrowleft} = (R^{\circlearrowleft} \circ L^{\circlearrowleft})^{-2} \circ F^{\circlearrowleft} \circ (R^{\circlearrowleft} \circ L^{\circlearrowleft})^{2}$$
$$T^{\circlearrowleft} = (R^{\circlearrowleft} \circ L^{\circlearrowleft})^{-3} \circ F^{\circlearrowleft} \circ (R^{\circlearrowleft} \circ L^{\circlearrowleft})^{3} \text{ or } (R^{\circlearrowleft} \circ L^{\circlearrowleft}) \circ F^{\circlearrowleft} \circ (R^{\circlearrowleft} \circ L^{\circlearrowleft})^{-1}.$$

The cublets are

$$\mathbf{1} := \{1, 5, 9\} \qquad \mathbf{2} := \{2, 13, 6\} \qquad \mathbf{3} := \{3, 17, 14\}$$
$$\mathbf{4} := \{4, 10, 18\} \qquad \mathbf{5} := \{15, 20, 24\} \qquad \mathbf{6} := \{11, 23, 19\}$$
$$\mathbf{7} := \{8, 22, 12\} \qquad \mathbf{8} := \{7, 16, 21\}.$$

We also consider $\underline{3}$–cycles in $\mathcal{S}_{24}$ for each cublet in the order they are written. We write the corresponding $\underline{3}$–cycles as $\mathbf{t_1}$, $\mathbf{t_2}$, etc. Therefore we have

$$\mathbf{t_1} := (1\ 5\ 9) \qquad \mathbf{t_2} := (2\ 13\ 6) \qquad \mathbf{t_3} := (3\ 17\ 14) \tag{1}$$
$$\mathbf{t_4} := (4\ 10\ 18) \qquad \mathbf{t_5} := (15\ 20\ 24) \qquad \mathbf{t_6} := (11\ 23\ 19) \tag{2}$$
$$\mathbf{t_7} := (8\ 22\ 12) \qquad \mathbf{t_8} := (7\ 16\ 21)\,. \tag{3}$$

Since these 8 $\underline{3}$–cycles are disjoint, they each commute with eachother. Therefore, if we define the permutation Σ as the product of the 8 $\mathbf{t}$'s in (1), it doesn't matter which order we write them.

Clearly $\mathcal{RK}_2$ preserves cubelets, but more is true. $\mathcal{RK}_2$ actually also preserves the cycles (and so $\mathcal{RK}_2 \leqslant C_{\mathcal{S}_{24}}(\Sigma)$.) This means if we draw an arrow from 1 to 5 to 9

This guarantees that $\mathcal{RK}_2$ does not contain a $\underline{2}$–cycle since no $\underline{2}$–cycle commutes Σ.

The homomorphism is given (on generators) as

$$\alpha\colon \mathcal{RK}_2 \to \mathcal{S}_8$$
$$: F^{\circlearrowleft} = (1\ 2\ 3\ 4)(5\ 13\ 17\ 10)(6\ 14\ 18\ 9) \quad \mapsto f^{\circlearrowleft} := (\mathbf{1\ 2\ 3\ 4})$$
$$: R^{\circlearrowleft} = (2\ 17\ 24\ 7)(3\ 20\ 21\ 6)(13\ 14\ 15\ 16) \mapsto r^{\circlearrowleft} := (\mathbf{2\ 3\ 5\ 8})$$
$$: L^{\circlearrowleft} = (1\ 18\ 23\ 8)(4\ 19\ 22\ 5)(9\ 10\ 11\ 12) \quad \mapsto l^{\circlearrowleft} := (\mathbf{1\ 4\ 6\ 7})\,.$$

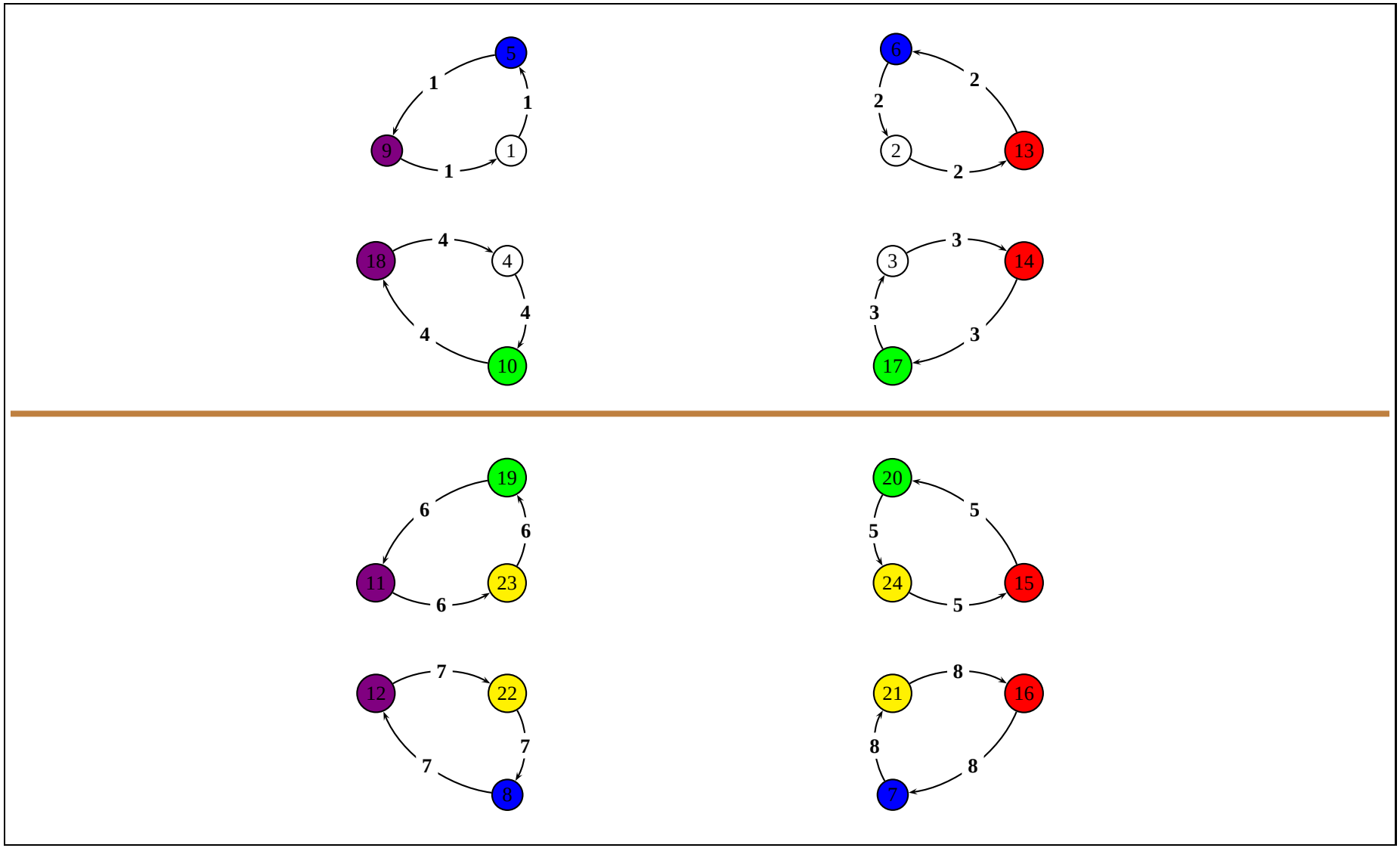
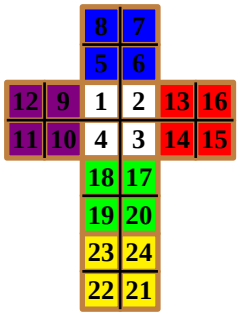
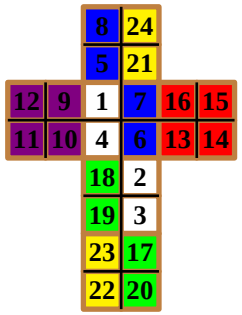


Figure 15.2: Grouped by cubelets

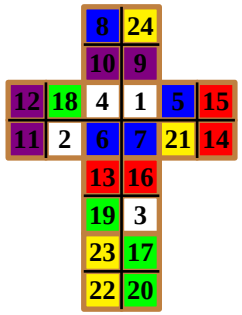
PROPOSITION 15.3. *Any permutation of cublets is possible. In other words the map α is surjective. Moreover the kernel is non-trivial, abelian and is contained inside of (a group isomorphic to) $\mathbb{Z}_3^8$.*

PROOF. Obviously the action of $\mathcal{RK}_2$ on the set of cubelets is 2-transitive. Therefore to show that α is surjective, we simply need to show that a 2–cycle is in the image (why?). But this just takes a bit of playing, I find

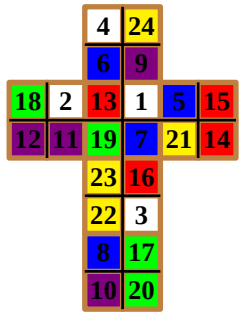




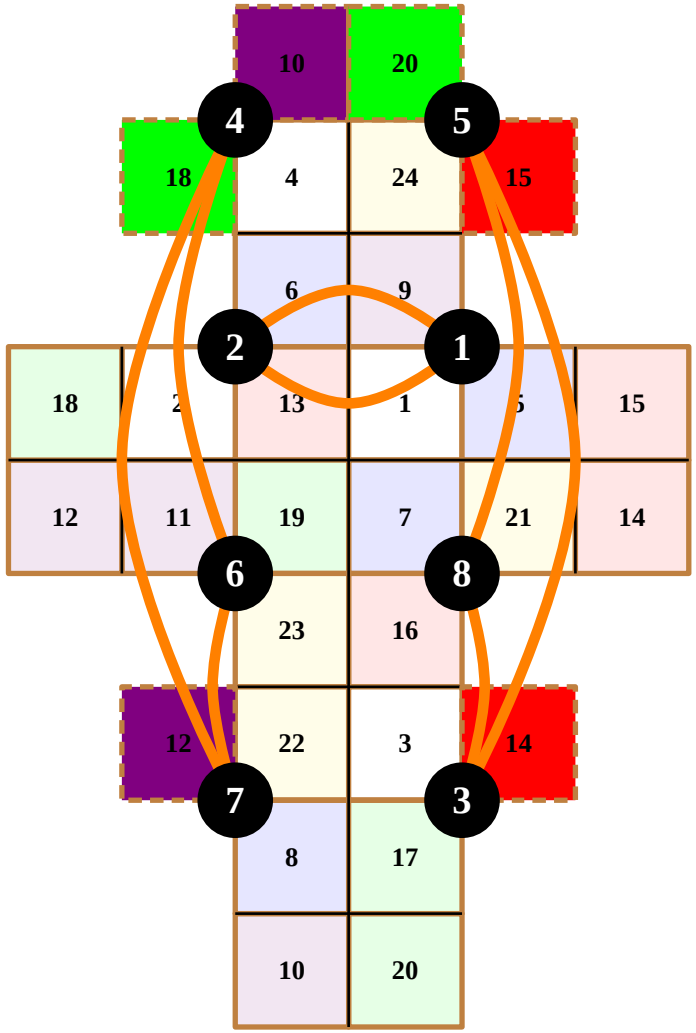
$R^{\circlearrowleft}$



$F^{\circlearrowleft} \circ R^{\circlearrowleft}$



$L^{\circlearrowleft -1} \circ F^{\circlearrowleft} \circ R^{\circlearrowleft}$



$$L^{\circlearrowleft -1} \circ F^{\circlearrowleft} \circ R^{\circlearrowleft} = (1\ 2\ 9\ 6\ 5\ 13)(3\ 20\ 21\ 14\ 15\ 16\ 17\ 24\ 7)(4\ 8\ 23\ 18\ 12\ 11\ 10\ 22\ 19)$$
$$\mapsto (\mathbf{1\ 2})(\mathbf{3\ 5\ 8})(\mathbf{4\ 7\ 6}).^{\text{I}}$$

(Therefore $L^{\circlearrowleft -1} \circ F^{\circlearrowleft} \circ R^{\circlearrowleft}$ acts as a $\underline{9}\,\underline{9}\,\underline{6}$–cycle in $\mathcal{RK}_2$ and a $\underline{3}\,\underline{3}\,\underline{2}$–cycle (via $l^{\circlearrowleft -1} \circ f^{\circlearrowleft} \circ r^{\circlearrowleft}$) on cublets. Therefore $\left(l^{\circlearrowleft -1} \circ f^{\circlearrowleft} \circ r^{\circlearrowleft}\right)^3$ is a $\underline{2}$–cycle in $\mathcal{S}_8$ and we are done. This also proves the kernel of α is non-trivial since we found an element of order $\text{lcm}(6, 9, 9) = 18$ which maps to an element of order $\text{lcm}(2, 3, 3) = 6$. {Finish} ■

Perhaps even more surprising is the fact that $\mathcal{RK}_2$ also does not contain a $\underline{3}$–cycle. In particular, none of the $\sigma_{\mathbf{i}}$'s are actually in $\mathcal{RK}_2$. As such $\mathcal{RK}_2 \leqslant C_{\mathcal{S}_{24}}(\Sigma)$ but $\mathcal{RK}_2 \neq C_{\mathcal{S}_{24}}(\Sigma)$. Indeed, clearly the only $\underline{3}$–cycles which are in the group $C_{\mathcal{S}_{24}}(\Sigma)$ are the $\mathbf{t}$'s from (1).

PROPOSITION 15.4. *$\mathcal{RK}_2$ does not contain a $\underline{3}$–cycle.*

PROOF. As pointed already, we only need to prove that $\mathcal{RK}_2$ does not contain any of the {Finish} ■

While $\text{Ker}(\alpha)$ has been proven to not contain a $\underline{3}$–cycle, the next lemma proves that the next thing "best thing" is true.

LEMMA 15.5.

This means that $\text{Ker}(\alpha)$ can be thought of as the subgroup of $\mathbb{Z}_3^8$ which are in the "hyper-plane"

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8.$$

In particular, this kernel has order $3^8/3 = 3^7$ (by **KIT**). Therefore we have proven

$$|\mathcal{RK}_2| = 3^7 \cdot 8! = 88,179,840$$

elements.

The more familiar $3 \times 3 \times 3$ Rubik's cube is now also not too much more difficult to crack.

SOLUTION

EXERCISE$^{\textbf{Hard!}}$ **15.3.** Prove that

$$|\mathcal{RK}_3| = (3^7 \cdot 8!) \cdot (12! \cdot 2^m), \quad \text{with } |\mathcal{RK}_2| = 3^7 \cdot 8!$$

for some integer m[II]

Hint! Focusing in on the 8 corners of the cube gives rise to a homomorphism $\mathcal{RK}_3 \to \mathcal{RK}_2$. The kernel of this homomorphism deals with the 24 new faces which come in pairs of two giving rise to 12 new cubelets (containing two faces per cubelet as opposed to 3) which correspond to each of the 12 edges of a cube.

ALGORITHM 15.6.

(1) Using moves which induce transpositions, move all cubelets back into their starting position.

(2) Now use moves in the abelian group $\text{Ker}(\alpha)$ to solve the puzzle.

[I] Colors only indicate faces in the same cubelet within a given cylce.

[II] Actually $m = 10$, but this is harder to prove by hand. Therefore the usual $3 \times 3 \times 3$ Rubik's cube group has a whopping

$$43,252,003,274,489,856,000$$

elements. This means, roughly, that if you sat down with an unsolved cube and made completely random sequences of moves at a rate of one thousand moves per second it would take you about 10 billion years to solve the puzzle. As compared to the one move per 10 years we'd have to make with a $2 \times 2 \times 2$ cube to finish in the same amount of time.

15.2 Final Questions

Exercise 15.4. Let G be the group

$$\langle (1\,4\,3\,9)(2\,7\,5\,6), (1\,3\,2\,6)(4\,5\,8\,7) \rangle .$$

(1) What is the order of G?
Hint! I choose a sucky generating set to ruin your day.

(2) Prove that G is centerless.

(3) Compute $[G, G]$ and G_{ab}.

(4) Show that G is solvable.

(5) For every p a prime divisor of $|G|$ write down a Sylow p–subgroup of G and calculate $n_p(G)$.

SOLUTION

IV APPENDICES

A Review of Cross Products and Dot Products

This Appendix reviews several basic ideas from any physics or multi-variable calculus class. The only section of this appendix which cannot be read before reading any other part of the book is the very first. For a reader who is looking at this appendix before looking at any other part of the book simply replace all instances of the symbol R with the symbol for the real numbers, $\mathbb{R}$. (Obviously you would need to ignore any explicit examples involving something besides the real numbers.)

A.1 Vector Arithmatic

Let R be a ring. The set of n-vectors in R, i.e. ordered n-tuples of elements in R

$$\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

can be added

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \text{ and } \mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \Rightarrow \mathbf{v} + \mathbf{w} = \begin{bmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{bmatrix}$$

and multiplied by a scalar, i.e. an element of $c \in R$,

$$c \cdot \mathbf{v} = c \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} cv_1 \\ \vdots \\ cv_n \end{bmatrix}.$$ [I]

In these situations there is a zero vector, $\mathbf{0}$, which is defined to be the vector whose every component is 0. That is

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Other important vectors are the *standard basis vectors*.

[I] They can also be multiplied by each other by multiplying together components, but this operation won't really be studied since, quite frankly, the multiplication sucks. For instance, it always introduces a ton of zero divisors: $(r, 0) \cdot (0, s) = (0, 0)$

DEFINITION A.1. The i^{th} *standard basis vector*, $\mathbf{e}_i$, is defined to be the vector [II]

$$\overset{i^{\text{th}}\text{-coord}}{\longrightarrow} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

whose only non-zero entry is a 1 in the i^{th} coordinate.

Thus in R^n, we have

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \qquad \cdots\cdots \qquad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

For example, in $\mathbb{R}^4$ we have

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

or written as rows,

$$\mathbf{e}_1 = (1,0,0,0) \quad \mathbf{e}_2 = (0,1,0,0) \quad \mathbf{e}_3 = (0,0,1,0), \quad \mathbf{e}_4 = (0,0,0,1).$$

In $\mathbb{Z}_6^3$, 1 is $[\,1\,]_6$, and we have

$$\mathbf{e}_1 = \begin{bmatrix} [\,1\,]_6 \\ [\,0\,]_6 \\ [\,0\,]_6 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} [\,0\,]_6 \\ [\,1\,]_6 \\ [\,0\,]_6 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} [\,0\,]_6 \\ [\,0\,]_6 \\ [\,1\,]_6 \end{bmatrix}.$$

The following properties come from applying the corresponding properties in R in a component by component case:

PROPOSITION A.2. *Let* $\mathrm{V} = R^n$. *Then, for any* $c, d \in R$ *and* $\mathbf{v}, \mathbf{w} \in \mathrm{V}$ *we have*

(1) $(\mathbf{v} + \mathbf{w}) + \mathbf{u} = \mathbf{v} + (\mathbf{w} + \mathbf{u})$

(2) $\mathbf{0} + \mathbf{v} = \mathbf{v} = \mathbf{v} + \mathbf{0}$

(3) $\mathbf{v} + (-1) \cdot \mathbf{v} = \mathbf{0}$.

(4) $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$

(5) $c \cdot (\mathbf{v} + \mathbf{w}) = c \cdot \mathbf{v} + c \cdot \mathbf{w}$

(6) $1 \cdot \mathbf{v} = \mathbf{v}$

(7) $c \cdot (d \cdot \mathbf{v}) = (cd) \cdot \mathbf{v}$

(8) $(c + d) \cdot \mathbf{v} = c \cdot \mathbf{v} + d \cdot \mathbf{v}$

Proof. Trivial. ∎

Part **(3)** means that it makes sense to define $(-1)\cdot\mathbf{v} = -\mathbf{v}$.

Rant A.3. If $R = \mathbb{R}$, the real numbers, in a Linear Algebra class these properties would be summed up simply by saying that V is a *vector space*. In such a class you learn that this is, in a sense, the only kind of vector space. This is true whenever R is a division ring, not just $\mathbb{R}$: for instance it is also true for R equal to $\mathbb{Q}$, $\mathbb{C}$, $\mathbb{H}$, or $\mathbb{Z}_p$ for p a prime. That is for these rings, any set V which satisfies the properties of the proposition and not too large (in the sense that is defined rigorously in an linear algebra class) for a given scalar multiplication is (essentially) R^n[III] for some positive integer n.

For other rings R, which are not fields, for example $R = \mathbb{Z}$ or $R = \mathbb{Z}_m$, where m is a composite integer, the properties of this proposition are summed up as saying that R^n is an *R-module*. R-modules are first encountered by most students in their first graduate algebra class. Even for a ring as "nice" as $\mathbb{Z}$, $\mathbb{Z}$-modules come in many more flavors than just $\mathbb{Z}^n$. Indeed any abelian group (first defined in Chapter 5) is a $\mathbb{Z}$ modules and of course many abelian groups are not, in any sense, $\mathbb{Z}^n$. The easiest such example is $\mathbb{Z}_2$!

A.2 Dot Products

Just like in a calculus class we will mostly be interested in the case that $R = \mathbb{R}$ and $n = 3$. It is good practice for a young budding mathematician such as yourself to look at each of the statements made below and ask yourself "is it really that important that I'm using $\mathbb{R}$ instead of say $\mathbb{C}$, $\mathbb{Q}$, or $\mathbb{Z}$?"

For $R \subseteq \mathbb{R}$ we can define the *length* of a vector $\mathbf{v}$, denoted $\|\mathbf{v}\|$, is defined as square root (assuming we are some place where this makes sense ... for instance, square roots do not really make sense over $\mathbb{C}$!) of the sum of the squares of its components. That is

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \Rightarrow \|\mathbf{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}.$$

It can be interrupted geometrically as the distance from the origin to the point (or endpoint) of $\mathbf{v}$. (This can be seen by using Pythogorean theorem.)

Note that for a scalar $c \in \mathbb{R}$ we have that

$$\|c\cdot\mathbf{v}\| = |c|\cdot\|\mathbf{v}\|$$

(where $|c|$ denotes the old fashioned absolute value of the real number c) since

$$\begin{aligned} \|c\cdot\mathbf{v}\| &= \sqrt{(cv_1)^2 + \cdots (cv_n)^2} \\ &= \sqrt{c^2\left(v_1^2 + \cdots v_n^2\right)} \\ &= \sqrt{c^2}\sqrt{v_1^2 + \cdots v_n^2} \\ &= |c|\cdot\|\mathbf{v}\|. \end{aligned}$$

[III] In a linear algebra class the statement is that every finite dimensional vector space is isomorphic to $\mathbb{R}^n$.

REMARK A.4. Since the square root does not always make sense, this definition of the length of a vector does not make sense (for say $R = \mathbb{C}$), the length squared,

$$v_1^2 + \cdots + v_n^2$$

does!

Given two n-vectors

$$\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \qquad \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

the dot product of $\mathbf{v}$ and $\mathbf{w}$ is defined to be the real number (yes I said number, not vector!)

$$\mathbf{v} \bullet \mathbf{w} \stackrel{\text{def}}{=} \sum_{i=1}^{n} v_i w_i.$$

Note that for any $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$ we have that

$$\mathbf{v} \bullet \mathbf{v} = v_1^2 + \cdots + v_n^2 = \|\mathbf{v}\|^2$$

the length squared. In particular $\mathbf{v} \bullet \mathbf{v} \geq 0$ with equality iff $\mathbf{v}$ is the zero vector $\mathbf{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$. We record this as a lemma for later.[I]

LEMMA A.5 (Non-Degeneracy). *Let* $\mathbf{v}$ *be a vector so that for all vectors* $\mathbf{u}$ *we have* $\mathbf{v} \bullet \mathbf{u} = 0$. *Then* $\mathbf{v} = \mathbf{0}$.

PROOF. The hypothesis guarantees us that $\mathbf{v} \bullet \mathbf{v} = 0$, but as just pointed out the only way that can happen is if $\mathbf{v} = \mathbf{0}$. ■

COROLLARY A.6. *Let* $\mathbf{v}$ *and* $\mathbf{w}$ *be vectors so that for all vectors* $\mathbf{u}$ *we have* $\mathbf{v} \bullet \mathbf{u} = \mathbf{w} \bullet \mathbf{u}$. *Then* $\mathbf{v} = \mathbf{w}$.

PROOF. If $\mathbf{v} \bullet \mathbf{u} = \mathbf{w} \bullet \mathbf{u}$, then $(\mathbf{v} - \mathbf{w}) \bullet \mathbf{u}$. Thus if the hypothesis holds, we have that $\mathbf{u} - \mathbf{w}$ has a zero dot product with all vectors. Hence, by the lemma, this mean that $\mathbf{v} - \mathbf{w} = \mathbf{0}$ or $\mathbf{v} = \mathbf{w}$. ■

EXERCISE A.1. (For those who already know about $\mathbb{Z}_n$ already.) The dot product is defined by the same formula for any ring. Find two examples (the example must include a ring R and a dimension n) where the dot product is not positive (i.e. there exists vectors $\mathbf{v}$ with $\mathbf{v} \bullet \mathbf{v} = 0$) but it is always non-degenerate.
Hint! For $n = 3$ and $R = \mathbb{Z}_3$, we have that $\mathbf{v} = (1, 1, 1)$ satisfies

$$\mathbf{v} \bullet \mathbf{v} = 1 + 1 + 1 = 3 \equiv_3 0.$$

However, we have that, say, $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1 \neq 0$.

SOLUTION

[I] Actually the lemma below is weaker than what we just stated. The goal of the mathematician is to look under the hood to figure out what is really going on. Therefore it is always a better bet to state and use the weakest possible propositions since they allow us to isolate what is really happening under the surface.

DEFINITION A.7. A function, B, with two vector inputs and one vector output is *bilinear* (or *doubly linear*) if for all scalars a, b, c, and d, and all vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{w}_1$, and $\mathbf{w}_2$ we have that

$$\mathrm{B}\Big((a\cdot\mathbf{v}_1+b\cdot\mathbf{v}_2),(c\cdot\mathbf{w}_1+d\cdot\mathbf{w}_2)\Big)=ac\cdot\mathrm{B}(\mathbf{v}_1,\mathbf{w}_1)+ad\cdot\mathrm{B}(\mathbf{v}_1,\mathbf{w}_2)+bc\cdot\mathrm{B}(\mathbf{v}_2,\mathbf{w}_1)+bd\cdot\mathrm{B}(\mathbf{v}_2,\mathbf{w}_2).$$

As written, the bilinear property is a bit hard to digest all at once. Perhaps the best way to think of it is as sort of a distributive property. Another way to think of the bilinear property is the think of it as a function is linear in each variable. In practice, the bilinear property is very easy to apply "one step at a time." For instance, for any bilinear function B we have that

$$\begin{aligned}\mathrm{B}(2\mathbf{e}_1+\mathbf{e}_2,\mathbf{e}_1+\mathbf{e}_2)&=\mathrm{B}(2\mathbf{e}_1,\mathbf{e}_1+\mathbf{e}_2)+\mathrm{B}(\mathbf{e}_2,\mathbf{e}_1+\mathbf{e}_2)\\&=2\mathrm{B}(\mathbf{e}_1,\mathbf{e}_1+\mathbf{e}_2)+\mathrm{B}(\mathbf{e}_2,\mathbf{e}_1+\mathbf{e}_2)\\&=2\mathrm{B}(\mathbf{e}_1,\mathbf{e}_1)+2\mathrm{B}(\mathbf{e}_1,\mathbf{e}_2)+\mathrm{B}(\mathbf{e}_2,\mathbf{e}_1+\mathbf{e}_2)\\&=2\mathrm{B}(\mathbf{e}_1,\mathbf{e}_1)+2\mathrm{B}(\mathbf{e}_1,\mathbf{e}_2)+\mathrm{B}(\mathbf{e}_2,\mathbf{e}_1)+\mathrm{B}(\mathbf{e}_2\mathbf{e}_2).\end{aligned}$$

With any further information we cannot simplify this any further. In Definition 3.197 on page 181 we go even further and define multilinear functions, i.e. functions of m-variables which are linear in each.

EXERCISE A.2. The bilinear property is equivalent to 3 properties in one! The three properties are

(Left Distributive)	$\mathrm{B}(\mathbf{v}_1+\mathbf{v}_2,\mathbf{w})=\mathrm{B}(\mathbf{v}_1,\mathbf{w})+\mathrm{B}(\mathbf{v}_2,\mathbf{w})$
(Right Distributive)	$\mathrm{B}(\mathbf{v}_1+\mathbf{v}_2,\mathbf{w})=\mathrm{B}(\mathbf{v}_1,\mathbf{w})+\mathrm{B}(\mathbf{v}_2,\mathbf{w})$
(Scalar Pulling)	$\mathrm{B}(c\cdot\mathbf{v},\mathbf{w})=c\cdot\mathrm{B}(\mathbf{v},\mathbf{w})=\mathrm{B}(\mathbf{v},c\cdot\mathbf{w})$.

Verify that these three properties are equivalent to the bilinear property just above.
Hint! To prove that the three properties imply the one simply apply the three properties in turn. To show that the one property implies the three choose some vectors to be zero and some scalars to be 1.

SOLUTION

EXERCISE A.3.

(1) Show that a bilinear function is determined by what it does to the standard basis (actually, if you've taken linear algebra you shouldn't have any more difficultly showing that this is the case for any basis) in the sense that if B is bilinear you can compute what $\mathrm{B}(\mathbf{v},\mathbf{w})$ is for any two vectors $\mathbf{v}$ and $\mathbf{w}$ if you know what each of the $\mathrm{B}(\mathbf{e}_i,\mathbf{e}_j)$ are.
Hint! Show for every

$$\mathbf{v}=c_1\mathbf{e}_1+\cdots c_n\mathbf{e}_n\quad\text{and}\quad\mathbf{w}=d_1\mathbf{e}_1+\cdots d_n\mathbf{e}_n$$

$(c_1,\ldots,c_n,d_1,\ldots,d_n\in\mathbb{R})$. Show that

$$\mathrm{B}(\mathbf{v},\mathbf{w})=\sum_{i,j=0}^{n}c_id_j\mathrm{B}(\mathbf{e}_i,\mathbf{e}_j).^{\text{II}}$$

SOLUTION

(2) For $n = 3$ suppose a bilinear function B satisfies $B(\mathbf{e}_a, \mathbf{e}_b) = a - b$ for any two standard basis vectors. Use the previous part to calculate

$$B(3\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + 5\mathbf{e}_3).$$

Here are some more properties of the dot product

THEOREM A.8. *The dot product is*

(1) **(Symmetric)** *For any two vectors* $\mathbf{v}$ *and* $\mathbf{w}$ *in* $\mathbb{R}^n$ *we have* $\mathbf{v} \bullet \mathbf{w} = \mathbf{w} \bullet \mathbf{v}$.

(2) **(Bilinear)** *Written out, this means for if for all scalars* $a, b, c,$ *and* $d \in \mathbb{R}$, *and all vectors* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_1,$ *and* $\mathbf{w}_2 \in \mathbb{R}^n$, *we have*

$$(a\mathbf{v}_1 + b\mathbf{v}_2) \bullet (c\mathbf{w}_1 + d\mathbf{w}_2) = ac(\mathbf{v}_1 \bullet \mathbf{w}_1) + ad(\mathbf{v}_1 \bullet \mathbf{w}_2) + bc(\mathbf{v}_2 \bullet \mathbf{w}_1) + bd(\mathbf{v}_2 \bullet \mathbf{w}_2).$$

PROOF. Since we can switch the order of the product of two numbers, the dot product is obviously symmetric. The bilinear property is based on the fact that multiplication of real numbers is bilinear, and the dot product is just the sum of such multiplications. ∎

EXAMPLE A.9. Let $\mathbf{v} = \begin{bmatrix}1\\2\\3\\2\end{bmatrix}$, $\mathbf{w} = \begin{bmatrix}0\\1\\0\\1\end{bmatrix}$, and $\mathbf{u} = \begin{bmatrix}-3\\-1\\1\\1\end{bmatrix}$. Then

$$\mathbf{v} \bullet \mathbf{v} = 1 + 4 + 9 + 4 = 18 \qquad \mathbf{v} \bullet \mathbf{w} = 0 + 2 + 0 + 2 = 4 \qquad \mathbf{v} \bullet \mathbf{u} = -3 - 2 + 3 + 2 = 0$$
$$* \qquad \mathbf{w} \bullet \mathbf{w} = 0 + 1 + 0 + 1 = 2 \qquad \mathbf{w} \bullet \mathbf{u} = 0 - 1 + 0 + 1 = 0$$
$$* \qquad * \qquad \mathbf{u} \bullet \mathbf{u} = 9 + 1 + 1 + 1 = 12$$

(We don't need to compute the places where a $*$ shows up since the dot product is symmetric.)

DEFINITION A.10.

(1) A vector $\mathbf{v}$ is said to be *normal* if $\mathbf{v} \bullet \mathbf{v} = 1$ (i.e. if it has length, $\|\mathbf{v}\|$, 1).

(2) The set of normal vectors in $\mathbb{R}^n$ is denoted $\mathbb{S}^n$ and called the *n-sphere*.

(3) A pair of vectors $\mathbf{v}$ and $\mathbf{w}$ are said to *orthogonal* if $\mathbf{v} \bullet \mathbf{w} = 0$. We denote this as $\mathbf{v} \perp \mathbf{w}$.

(4) The set of vectors orthogonal to a given vector $\mathbf{v}$, is denoted $\mathbf{v}^\perp$.

[II] The fact should be more intuitive than its proof. The idea is that the bilinear property allows you to "break things down" enough so you simply need to compute B applied to pairs of standard basis vectors.

(5) A set of n-vectors

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$$

are said to be *orthonormal* if for all $\mathbf{v}_i$ and $\mathbf{v}_j$ ($i, j \in \overline{m}$) we have that

$$\mathbf{v}_i \bullet \mathbf{v}_j = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise.} \end{cases} .$$

If addition $m = n$ we say that the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is an *orthonormal basis*.

In other words a set of vectors are orthonormal iff each vector is normal and each pair of distinct vectors in the set are orthogonal. For instance the vectors $(1, 2, 3)$ and $(-2, 1, 0)$ are orthogonal and the vector $\mathbf{u}$ from Example A.9 is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$ from that same example.

Remark A.11. A basic property of the dot product is that it measures the angle between two non-zero vectors. This can be seen by using the law of cosines which says for a triangle with vertices $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{u}$ we have that $c^2 = a^2 + b^2 - 2ab\cos\theta$ where θ is the angle between the side of length a and the side of length b.[III] Now using bilinearity and the fact that for any vector $\mathbf{u}$ we have that $\mathbf{u} \bullet \mathbf{u} = \|\mathbf{u}\|^2$ we see

$$\begin{aligned} \|\mathbf{v} - \mathbf{w}\|^2 &= (\mathbf{v} - \mathbf{w}) \bullet (\mathbf{v} - \mathbf{w}) \\ &= \mathbf{v} \bullet \mathbf{v} - 2\mathbf{v} \bullet \mathbf{w} + \mathbf{w} \bullet \mathbf{w} \\ &= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2 \cdot (\mathbf{v} \bullet \mathbf{w}). \end{aligned}$$

Using this on a triangle with a vertices at the origin, at $\mathbf{v}$, and at $\mathbf{w}$ we have that the side lengths of this triangle are $\|\mathbf{v}\|$, $\|\mathbf{w}\|$ and $\|\mathbf{v} - \mathbf{w}\|$ respectively, so comparing this with the law of cosines yields the famous formula

$$\frac{\mathbf{v} \bullet \mathbf{w}}{\sqrt{\mathbf{v} \bullet \mathbf{v}} \cdot \sqrt{\mathbf{w} \bullet \mathbf{w}}} = \frac{\mathbf{v} \bullet \mathbf{w}}{\|\mathbf{v}\| \cdot \|\mathbf{w}\|} = \cos\theta_{\mathbf{v},\mathbf{w}}$$

(where $\theta_{\mathbf{u},\mathbf{v}}$ is the (unorientated) angle[IV] (restricted to be between 0 and π) between the arrows connecting the origin to the points $\mathbf{v}$ and $\mathbf{w}$.) This is the most efficient way of calculating the angle between two vectors. In particular vectors are orthogonal to each other iff $\theta_{\mathbf{v},\mathbf{w}} = \frac{\pi}{2}$, i.e. iff they perpendicular to each other.

Exercise A.4. Define

$$\mathbf{v} = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix}, \quad \mathbf{w} = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \quad \mathbf{u} = \frac{1}{3}\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}.$$

Verify that $\{\mathbf{v}, \mathbf{u}, \mathbf{w}\}$ is an orthonormal basis.

SOLUTION

[III] This is proved very easily by dropping a perpendicular and on the triangle on the right which has lengths $a - \cos\theta\, b$, c, and $b \sin\theta$.

[IV] So $\theta_{\mathbf{u},\mathbf{v}} = \theta_{\mathbf{v},\mathbf{u}}$.

In terms of linear algebra, this exercise says that $\mathbf{v}^\perp$ is a subspace.

EXERCISE A.5. Suppose that $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{v}^\perp$ and $c \in \mathbb{R}$, then prove that $\mathbf{w}_1 + \mathbf{w}_2$ and $c\mathbf{w}_1 \in \mathbf{v}^\perp$.

SOLUTION

EXERCISE A.6. Suppose that for all vectors $\mathbf{u}$ we have that $\mathbf{v} \bullet \mathbf{u} = \mathbf{w} \bullet \mathbf{u}$. Then show that $\mathbf{v} = \mathbf{w}$.
Hint! Apply Lemma A.5 to the vector $\mathbf{v} - \mathbf{w}$.

SOLUTION

Notice that in any dimension n we always have that

$$\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$$

forms an orthonormal basis.

In $\mathbb{R}^3$ the standard basis vectors get special names:

$$\mathbf{i} \stackrel{\text{def}}{=} \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{j} \stackrel{\text{def}}{=} \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \text{and} \qquad \mathbf{k} \stackrel{\text{def}}{=} \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

We will later see the historical significance of these names. These vectors are called i-hat (*eye hat*), j-hat (*jay hat*), and k-hat (*kay hat*) respectively. This is because when they are written by hand they are written as

$$\hat{\imath}, \hat{\jmath}, \text{ and } \hat{k}.$$

Of course this difference is because people cannot write in bold and unbold by hand! Also, putting hats over things too often in type will eventually look horrible. Table A.1 records the dot product of these standard basis vectors in $\mathbb{R}^3$.

•	**i**	**j**	**k**
i	1	0	0
j	0	1	0
k	0	0	1

Table A.1: The dot product of **i**, **j**, and **k**.

A.3 CROSS PRODUCT

In calculus 3, linear algebra, or elementary physics students learn about the cross product of two vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$, denoted $\mathbf{v} \times \mathbf{w}$. Until stated otherwise all vectors are 3-vectors. (If you jump ahead to the section on determinants you can get the formula for the determinant.)

DEFINITION A.12. The *cross product* of two vectors $\mathbf{v}$ and $\mathbf{w}$ is the unique vector, $\mathbf{v} \times \mathbf{w}$, so that for any third vector $\mathbf{u}$ we have

$$(\mathbf{v} \times \mathbf{w}) \bullet \mathbf{u} = \det(\mathbf{v}, \mathbf{w}, \mathbf{u}).$$

It should be pointed out that while the dot product was a rule assigning two vectors (defined in any dimension) a real number (i.e., for any positive integer n it is a function $\bullet : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$), the cross product only makes sense in 3 dimensions and is a rule which assigns any two 3–vectors another 3–vector, (i.e. it is a function $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$). For this reason the cross product is often called a *vector product* while the dot product is often called a *scalar product*. Notice that since

$$\det(\mathbf{v}, \mathbf{w}, \mathbf{v}) = \det(\mathbf{v}, \mathbf{w}, \mathbf{w}) = 0 \Rightarrow (\mathbf{v} \times \mathbf{w}) \perp \mathbf{v}, \mathbf{w}.$$

In words the cross product of two vectors is orthogonal to both of the original vectors. Moreover, whenever the cross product is non-zero[I] we have

$$0 < (\mathbf{u} \times \mathbf{v}) \bullet (\mathbf{u} \times \mathbf{v}) = \det(\mathbf{u}, \mathbf{v}, \mathbf{u} \times \mathbf{v})$$

and so the triple $\mathbf{u}, \mathbf{v}$, and $\mathbf{u} \times \mathbf{v}$ is right-handed.

Let's show how this seemingly strange and abstract definition can be used to make actual computations:

EXAMPLE A.13. I hope this example convinces you that this is well defined. Let $\mathbf{u} = \mathbf{i} + \mathbf{j}$ and $\mathbf{v} = \mathbf{j} + \mathbf{k}$. Then

$$\mathbf{u} \times \mathbf{v} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$$

for some real numbers a, b, and c. We need to solve for these real numbers. To do this let's take the dot product of both sides with the vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$.

$$\begin{aligned}
a &= (a\mathbf{i} + b\mathbf{j} + c\mathbf{k}) \bullet \mathbf{i} = (\mathbf{u} \times \mathbf{v}) \bullet \mathbf{i} = \det(\mathbf{i} + \mathbf{j}, \mathbf{j} + \mathbf{k}, \mathbf{i}) = \det(\mathbf{j}, \mathbf{j} + \mathbf{k}, \mathbf{i}) = \det(\mathbf{j}, \mathbf{k}, \mathbf{i}) = +1. \\
b &= (a\mathbf{i} + b\mathbf{j} + c\mathbf{k}) \bullet \mathbf{j} = (\mathbf{u} \times \mathbf{v}) \bullet \mathbf{j} = \det(\mathbf{i} + \mathbf{j}, \mathbf{j} + \mathbf{k}, \mathbf{j}) = \det(\mathbf{i}, \mathbf{j} + \mathbf{k}, \mathbf{j}) = \det(\mathbf{i}, \mathbf{k}, \mathbf{j}) = -1. \\
c &= (a\mathbf{i} + b\mathbf{j} + c\mathbf{k}) \bullet \mathbf{k} = (\mathbf{u} \times \mathbf{v}) \bullet \mathbf{k} = \det(\mathbf{i} + \mathbf{j}, \mathbf{j} + \mathbf{k}, \mathbf{k}) = \det(\mathbf{i} + \mathbf{j}, \mathbf{j}, \mathbf{k}) = \det(\mathbf{i}, \mathbf{j}, \mathbf{k}) = +1.
\end{aligned}$$

Therefore $\mathbf{u} \times \mathbf{v} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k} = \mathbf{i} - \mathbf{j} + \mathbf{k}$.

THEOREM A.14. *The cross product is*

(1) (Skew-Symmetric) $\mathbf{v} \times \mathbf{v} = \mathbf{0}$

(2) (Bilinear [or Doubly Linear]) *Written out, this means for any four vectors* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_1$ *and* $\mathbf{w}_2$ *in* $\mathbb{R}^3$ *and scalars* a, b, c, *and* $d \in \mathbb{R}$ *we have*

$$\begin{aligned}
(a\mathbf{v}_1 + b\mathbf{v}_2) \times (c\mathbf{w}_1 + d\mathbf{w}_2) = {} & ac(\mathbf{v}_1 \times \mathbf{w}_1) + ad(\mathbf{v}_1 \times \mathbf{w}_2) \\
& + bc(\mathbf{v}_2 \times \mathbf{w}_1) + bd(\mathbf{v}_2 \times \mathbf{w}_2).
\end{aligned}$$

PROOF.

(1) Let $\mathbf{v} \in \mathbb{R}^3$. Then for any vector $\mathbf{u}$ we have that

$$(\mathbf{v} \times \mathbf{v}) \bullet \mathbf{u} = \det(\mathbf{v}, \mathbf{v}, \mathbf{u}) = 0.$$

But by Lemma A.5 this implies that $\mathbf{v} \times \mathbf{v} = \mathbf{0}$.

[I] We will see this is the case whenever $\mathbf{v}$ and $\mathbf{w}$ are not parallel

(2) We just show that for any vectors $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{w}$ we have that

$$(\mathbf{v}_1 + \mathbf{v}_2) \times \mathbf{w} = \mathbf{v}_1 \times \mathbf{w} + \mathbf{v}_2 \times \mathbf{w}$$

and leave the rest to you. By Exercise A.6 it is enough to show that they have the same dot product with a arbitrary vector $\mathbf{u}$. That is our goal is to show that

$$((\mathbf{v}_1 + \mathbf{v}_2) \times \mathbf{w}) \bullet \mathbf{u} = (\mathbf{v}_1 \times \mathbf{w} + \mathbf{v}_2 \times \mathbf{w}) \bullet \mathbf{u}$$

for the arbitrary vector $\mathbf{u}$. Indeed

$$\begin{aligned} ((\mathbf{v}_1 + \mathbf{v}_2) \times \mathbf{w}) \bullet \mathbf{u} &= \det(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{w}, \mathbf{u}) \\ &= \det(\mathbf{v}_1, \mathbf{w}, \mathbf{u}) + \det(\mathbf{v}_2, \mathbf{w}, \mathbf{u}) \\ &= (\mathbf{v}_1 \times \mathbf{w}) \bullet \mathbf{u} + (\mathbf{v}_2 \times \mathbf{w}) \bullet \mathbf{u} \\ &= (\mathbf{v}_1 \times \mathbf{w} + \mathbf{v}_2 \times \mathbf{w}) \bullet \mathbf{u} \end{aligned}$$

as desired. ■

The following exercise shows that two properties could be called the skew–symmetric (or alternating) property.

EXERCISE A.7. Let $\mathcal{A}: \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ be a bilinear function[II]. Prove that (1) ⇔ (2) where

(1) For all $\mathbf{V}$ in $\mathbb{R}^3$

$$\mathcal{A}(\mathbf{V}, \mathbf{V}) = \mathbf{0}.$$

(2) For all $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$

$$\mathcal{A}(\mathbf{v}, \mathbf{w}) = -\mathcal{A}(\mathbf{w}, \mathbf{v})$$

Hint! Expand $\mathbf{0} = \mathcal{A}((\mathbf{v} + \mathbf{w}), (\mathbf{v} + \mathbf{w}))$ and use **(1)** with $\mathbf{V} = \mathbf{v} + \mathbf{w}$ to see that **(1)⇒ (2)**.

SOLUTION

Using the basic properties of the determinant, that for any permutation σ,

$$\det(\mathbf{e}_{\sigma(1)}, \ldots, \mathbf{e}_{\sigma(n)}) = \varepsilon(\sigma)$$

we can complete a table of what the cross product of any two standard basis vectors is:[III]

[II] It doesn't really matter that we are dealing with $\mathbb{R}^3$ here. The proof you give will (well it should) apply to any bilinear function

$$R^n \times R^m \to R^s$$

(n, m, and $s \in \mathbb{Z}^+$) where R is any ring where 2 is not a zero divisor. (Your proof will not work $\mathbb{Z}_2$ but it'll work for all other fields we consider in the main part of the text.)

[III] Tables like these are always read as the row entry × the column entry.

×	**i**	**j**	**k**
i	**0**	**k**	**−j**
j	**−k**	**0**	**i**
k	**j**	**−i**	**0**

Table A.2: The cross product of **i**, **j**, and **k**.

EXAMPLE A.15. We use just the properties from Theorem A.14 and Table A.2 to make the following computation:

$$\begin{aligned}
&(\mathbf{i}+\mathbf{j}+3\mathbf{k})\times(\mathbf{i}+\mathbf{j}-\mathbf{k})\\
&\quad=(\mathbf{i}+\mathbf{j})\times(\mathbf{i}+\mathbf{j}-\mathbf{k})+3\mathbf{k}\times(\mathbf{i}+\mathbf{j}-\mathbf{k})\\
&\quad=(\mathbf{i}+\mathbf{j})\times(\mathbf{i}+\mathbf{j})+(\mathbf{i}+\mathbf{j})\times(-\mathbf{k})+3\mathbf{k}\times(\mathbf{i}+\mathbf{j})+3\mathbf{k}\times(-\mathbf{k})\\
&\quad=(\mathbf{i}+\mathbf{j})\times(\mathbf{i}+\mathbf{j})-(\mathbf{i}+\mathbf{j})\times\mathbf{k}+3(\mathbf{k}\times(\mathbf{i}+\mathbf{j}))-3(\mathbf{k}\times\mathbf{k})\\
&\quad=\mathbf{0}-(\mathbf{i}+\mathbf{j})\times\mathbf{k}+3(\mathbf{k}\times(\mathbf{i}+\mathbf{j}))+\mathbf{0}\\
&\quad=-(\mathbf{i}+\mathbf{j})\times\mathbf{k}+3\mathbf{k}\times(\mathbf{i}+\mathbf{j})\\
&\quad=-(\mathbf{i}\times\mathbf{k}+\mathbf{j}\times\mathbf{k})+3(\mathbf{k}\times\mathbf{i}+\mathbf{k}\times\mathbf{j})\\
&\quad=-(-\mathbf{j}+\mathbf{i})+3(\mathbf{j}-\mathbf{i})\\
&\quad=\mathbf{j}-\mathbf{i}+3\mathbf{j}-3\mathbf{i}\\
&\quad=-4\mathbf{i}+4\mathbf{j}.
\end{aligned}$$

EXERCISE A.8. For any two vectors

$$\mathbf{v}=v_1\mathbf{i}+v_2\mathbf{j}+v_3\mathbf{k} \text{ and } \mathbf{w}=w_1\mathbf{i}+w_2\mathbf{j}+w_3\mathbf{k}$$

use just the properties from Theorem A.14 and Table A.2 to derive the following general formula for the cross product $\mathbf{v}\times\mathbf{w}$,

$$(v_2w_3-v_3w_3)\mathbf{i}+(v_3w_1-v_1w_3)\mathbf{j}+(v_1w_2-w_1v_2)\mathbf{k}.^{IV}$$

Conclude that those properties uniquely determine the cross product (they led to a formula!).

SOLUTION

For the following lemma and proposition we assume some basic properties of the determinant and the transpose. Note, these formulas could be proved via a long a slog by directly computing with the formula. This is the approach of most elementary texts on the subject, and if you're interested you can consult such a text. We give a proof that involves minimal calculation.

LEMMA A.16. *For any two* 3*-vectors,* $\mathbf{v}$ *and* $\mathbf{w}$*, we have*

$$\|\mathbf{v}\times\mathbf{w}\|=\sqrt{\|\mathbf{v}\|^2\cdot\|\mathbf{w}\|^2-(\mathbf{v}\bullet\mathbf{w})^2}.$$

PROOF. If $\mathbf{v}$ and $\mathbf{w}$ are proportional, then $\|\mathbf{v}\times\mathbf{w}\|=0$ and the formula checks. Otherwise $\|\mathbf{v}\times\mathbf{w}\|\neq 0$ and we proceed with this case.

As with most proofs involving length, we prove that the square of the formula holds. For starters, note that we have

$$\begin{aligned} \|\mathbf{v}\times\mathbf{w}\|^2 &= (\mathbf{v}\times\mathbf{w})\bullet(\mathbf{v}\times\mathbf{w}) \\ &= \det(\mathbf{v},\mathbf{w},\mathbf{v}\times\mathbf{w}). \end{aligned}$$

Now define the matrix $\mathbf{A}$ to be the matrix under the determinant, i.e. the matrix $\left(\mathbf{v}\mid\mathbf{w}\mid\mathbf{v}\times\mathbf{w}\right)$. The calculation we just did proves that

$$\det\mathbf{A} = \|\mathbf{v}\times\mathbf{w}\|^2 .$$

Now

$$\mathbf{A}^\top\cdot\mathbf{A} = \begin{pmatrix} \mathbf{v}\bullet\mathbf{v} & \mathbf{v}\bullet\mathbf{w} & \mathbf{v}\bullet(\mathbf{v}\times\mathbf{w}) \\ \mathbf{w}\bullet\mathbf{v} & \mathbf{w}\bullet\mathbf{w} & \mathbf{w}\bullet(\mathbf{v}\times\mathbf{w}) \\ (\mathbf{v}\times\mathbf{w})\bullet(\mathbf{v}\times\mathbf{w}) & (\mathbf{v}\times\mathbf{w})\bullet\mathbf{w} & (\mathbf{v}\times\mathbf{w})\bullet(\mathbf{v}\times\mathbf{w}) \end{pmatrix} = \begin{pmatrix} \|\mathbf{v}\|^2 & \mathbf{v}\bullet\mathbf{w} & 0 \\ \mathbf{v}\bullet\mathbf{w} & \|\mathbf{w}\|^2 & 0 \\ 0 & 0 & \|\mathbf{v}\times\mathbf{w}\|^2 \end{pmatrix}.$$

See Remark or simply do the simple calculation out to see the first equality. The second equality comes from using the fact that the cross product of any two vectors is always perpendicular to either vector involved.

Therefore,

$$\det\left(\mathbf{A}^\top\cdot\mathbf{A}\right) = \left(\|\mathbf{v}\|^2\cdot\|\mathbf{w}\|^2 - (\mathbf{v}\bullet\mathbf{w})^2\right)\cdot\|\mathbf{v}\times\mathbf{w}\|^2 . \tag{1}$$

But we also have that

$$\det\left(\mathbf{A}^\top\cdot\mathbf{A}\right) = \det\mathbf{A}^\top\cdot\det\mathbf{A} = \det\mathbf{A}\cdot\det\mathbf{A} = \det(\mathbf{A})^2 = \|\mathbf{v}\times\mathbf{w}\|^4 . \tag{2}$$

Comparing these two equations, (1) and (2), and canceling the non-zero $\|\mathbf{v}\times\mathbf{w}\|^2$ proves the formula. ■

PROPOSITION A.17. $\dfrac{\|\mathbf{v}\times\mathbf{w}\|}{\|\mathbf{v}\|\cdot\|\mathbf{w}\|} = \sin\theta_{\mathbf{v},\mathbf{w}}$

PROOF. Remark A.11 says that

$$\frac{\mathbf{v}\bullet\mathbf{w}}{\|\mathbf{v}\|\cdot\|\mathbf{w}\|} = \cos\theta_{\mathbf{v},\mathbf{w}} \Rightarrow \mathbf{v}\bullet\mathbf{w} = \|\mathbf{v}\|\cdot\|\mathbf{w}\|\cos\theta_{\mathbf{v},\mathbf{w}}.$$

Plugging this into the formula from the lemma, Lemma A.16, shows us that

$$\begin{aligned} \|\mathbf{v}\times\mathbf{w}\|^2 &= \|\mathbf{v}\|^2\cdot\|\mathbf{w}\|^2 - (\mathbf{v}\bullet\mathbf{w})^2 \\ &= \|\mathbf{v}\|^2\cdot\|\mathbf{w}\|^2 - \|\mathbf{v}\|^2\cdot\|\mathbf{w}\|^2\cos^2\theta_{\mathbf{v},\mathbf{w}} \\ &= \|\mathbf{v}\|^2\cdot\|\mathbf{w}\|^2(1-\cos^2\theta_{\mathbf{v},\mathbf{w}}) \\ &= \|\mathbf{v}\|^2\cdot\|\mathbf{w}\|^2\sin^2\theta_{\mathbf{v},\mathbf{w}}. \end{aligned}$$

Taking square roots and paying attention to the sign of the sin proves the results. ■

B General Associativity and Cayley's theorem

In Chapter 4 we have already discussed generalized associativity. In fact, we stated, without proof, Theorem 4.20. We will now see a proof of this theorem along the way of proving another very important theorem, Cayley's Theorem.[1] This appendix is best read concurrently with Chapter 5

Let's begin by discussing the problem again. The definition of a group only requires that we can reassociate triples of elements. The axiom reads:

$$\forall a, b, c \in G, (ab)c = a(bc).$$

It does say anything about higher order products. The question is, can we always ignore parenthesis? Fortunately, the answer is yes. To see how little fun proving this can be, let's show that $((ab)(cd))((xy)z)) = a(b(c(d(x(yz)))))$.

$$\begin{aligned}((ab)(cd))((xy)z)) &= (ab)((cd))((xy)z))) \\ &= a(b((cd)((xy)z))) \\ &= a(b(cd))((xy)z))) \\ &= a(b(cd((xy)z)))) \\ \text{AHHHHHHHH} & \\ &= a(b(c(d((x(yz)))))).\end{aligned}$$

This certainly isn't right ... the "AHHH" line in the middle is extremely suspect, and I got tired of counting parenthesis. For a boring, tedious proof along these line, of the more general result check out the book. It should be intuitive to proceed along these lines, we'd have to do some kind of induction, but induction over what? We take a different, more abstract approach. This more abstract approach, in addition to kicking butt, allows us to prove something that we will want to prove eventually anyway: Cayley's theorem.

Here's the idea of the proof. If a, b, c, d, x, y and z were functions, then the proof that both ways of reassociating them if we apply both groupings of functions to a point (quick, what symbol haven't I already used???) p we would get $a(b(c(d(x(y(z(p)))))))$ in either case (and in any way we stick the parenthesis). Since functions are determined by what they do to points, these two functions would be the same. Of course the sentence *a, b, c ... were functions* is akin to saying if I could fly like superman I would go to Mars for Easter break. We are talking about an abstract group! Elements of an abstract group are, well, abstract ... and functions are much more (maybe not to a calc I student ...) concrete.

We can however stick a group inside a set of functions in a way that the multiplication in G "becomes" composition. For a set X denote the set of all functions from X to itself by $\mathcal{F}_X$. As already pointed out, $(\mathcal{F}_X, \circ)$ is a monoid with identity given by the identity function. And this monoid doesn't just satisfy the associative property, it satisfies the generalized associativity property (even composition of relations satisfy this, see Exercise 2.18 on page 22). The case we are interested in is the case that X is the set G itself. We are going to define an injective function $\Phi\colon G \to \mathcal{F}_X$ which satisfies $\Phi(a \cdot b) = \Phi(a) \circ \Phi(b)$ for all a and b in G (I brought the $\cdot$ back for emphasis). How does this help us? Well since in the right hand side we don't need parenthesis, and since the function Φ is injective, that means we won't need parenthesis either in G. This is because any way of associating a product in G will go to a way off associating a single product in $\mathcal{F}_G$. But the set $\mathcal{F}_G$ is a set of functions so there is only one way of associating

[1] We will later give a much faster proof of Cayley's theorem within the more general context of group actions, see Theorem 10.13 coming up on page 486.

a product. Thus any two ways of associating a product in G map, under and injective map, to a **single** element. Therefore, they must have been the same in the first place. What's more is that it will be clear that not only does the image of Φ consist of functions, it will consist of bijections. So if $|G| = n$, this gives us a way of thinking of ANY (ok ... finite, but it also works for infinite groups, but who cares about infinite permutation groups??) group G as a subset of $\mathcal{S}_n$.

Ok, enough blabbing ... what could Φ possibly be? What natural function can we associate to every element of G? How about left multiplication by g? (Right multiplication would "almost" work too, but things would get a little mixed up). That is we define for every $g \in G$ we define a function $\Phi(g)$ via

$$\begin{aligned}\Phi(g)\colon G &\to G\\ : a &\mapsto g \cdot a.\end{aligned}$$

You might be thinking a function to functions??? Then you might start to spark and shut down. But of course this is something you should be familiar with from calculus classes! When you have a function of two variables you can always fix one of the variables and let the other one vary. This means you are assigning a point (the one you fixed) to a function ... which is determined by the other variable. For instance the function $\sin(xy)$ can be thought of as written (two inputs, one output) or as first fixing x and letting y vary. So it sends the number 2, to the function $\sin 2y$, the number 17 to the function $\sin 17y$ and so on. We got this function Φ by fixing one "variable" in the function

$$\cdot\colon G \times G \to G,$$

i.e. (fixing g and letting a vary) $\Phi(g)(a) = g \cdot a$.

Back to proving generalized associativity: I made two promises about this dude Φ:

PROPOSITION B.1.

(1) *The assignment $g \mapsto \Phi(g)$ is one-to-one (i.e. if $g \neq g' \Rightarrow \Phi(g) \neq \Phi(g')$).*

(2) *For all $a, b \in G$ we have $\Phi(a \cdot b) = \Phi(a) \circ \Phi(b)$.*

PROOF.

(1) We need to show that if $\Phi(g_1)$ and $\Phi(g_2)$ are the same functions, that the elements g_1 and g_2 are themselves equal. We are lucky that the requirement that two functions be equal is a strong one, in our case we know that for all $g \in G$ we must have

$$\Phi(g_1)(g) = \Phi(g_2)(g).$$

So we have our pick of what to use for g, let's take the only element we know for sure to be in G, the identity, $e = e_G$. If $\Phi(g_1) = \Phi(g_2)$ are equal functions, then they must do the same thing to e_G. What does $\Phi(g_1)$ do to e_G? $\Phi(g_1)$ sends e to $g_1 \cdot e_G = g_1$. At the same time $\Phi(g_2)$ sends to e to $g_2 \cdot e = g_2$. So in particular, we must have $g_1 = g_2$, which is all we really needed!

(2) Let a and $b \in G$ be arbitrary. To show the equality of two functions, $\Phi(a \cdot b)$ and $\Phi(a) \circ \Phi(b)$, we need to show that they do exactly the same thing to every point x in the domain, which here is G. What does $\Phi(a \cdot b)$ do to x? Well, it takes it to $(a \cdot b) \cdot x$. What does $\Phi(a) \circ \Phi(b)$ do to x? Well,

$$\begin{aligned}(\Phi(a) \circ \Phi(b))\, x &= \Phi(a)\,(\Phi(b)(x)) && \text{(Def'n of comp of functions)}\\ &= \Phi(a)(b \cdot x) && \text{(Def'n of } \Phi(b)\text{)}\\ &= a \cdot (b \cdot x) && \text{(Def'n of } \Phi(a)\text{).}\end{aligned}$$

Thus the first dude, $\Phi(a \cdot b)$ takes x to $(a \cdot b) \cdot x$ and the second dude takes x to $a \cdot (b \cdot x)$! These are equal exactly, and precisely because of the associative property!! ■

COROLLARY B.2 (Generalized Associativity)**.** (*Same as Theorem 4.20*) *In an associative algebraic system you can re-associate any product of a finite string of elements.*[II]

[II] See remark below to see why we have really proven this statement for any semi-group.

Proof. Fix an integer n. Define a relation $\mathcal{R}$ on G^n via $\mathcal{R}((g_1, \ldots, g_n), g)$ if we can re-associate the product

$$g_1(g_2(g_3(\cdots(g_{n_1}g_n))))$$

so that it equals g. Notice that the theorem holds iff $\mathcal{R}$ is single valued (and is hence a function). But, using the above and the fact that functions satisfy generalized associativity, we know that the only thing that $(g_1, \ldots, g_n)$ can be related to is

$$(\Phi(g_1) \circ \cdots \circ \Phi(g_n))(e_G).$$

■

Remark B.3. We never used the fact that any element of G has an inverse. Therefore, this proof works just as well for any monoid, thus generalized associativity holds for any monoid. What if simply have an associative binary operator on a set M (i.e. throw away the condition that M have an identity)? We certainly used the fact that there was an identity in G in the proof. But there's a bit of a trick here to make it work even in this case. We can extend M in order to force it to have an identity and the same proof carries over. For instance the set of even integers form are associative algebraic system under multiplication but they lack an identity (who doesn't now a days?). However, if we consider the set of even integers **and** 1, then this new set is a monoid (of course you can't add in this new set, but monoids only require one operation).

We are therefore done proving generalized associativity for groups. But while we are on the subject, we might as well finish the proof of Cayley's theorem.

Theorem B.4 (Cayley). *Every group G has the same multiplication table as a subset of $\mathcal{S}_G$.(Very roughly, this says that every finite group is a subgroup of some permutation group $\mathcal{S}_n$.)*

Proof. The only thing we need to show is that the image of Φ is contained in the bijections of G (and not just functions).[III] That is we need to show that for every $g \in G$, that $\Phi(g)$ is a bijection. As is the case with proving that anything is a bijection, we could either show that $\Phi(g)$ is one-to-one and onto for any g or we could just find any inverse. We go the latter route (but you should note that the fact that the $\Phi(g)$'s are one-to-one is simply a restatement of the left cancelation law, surjectivity is proven by finding an inverse anyway ...). I claim that the inverse of $\Phi(g)$ is $\Phi(g^{-1})$ (how else could we undo multiplying on the left by g??). We wanna show that

$$\Phi(g) \circ \Phi(g^{-1}) = \mathbf{1}_G = \Phi(g) \circ \Phi(g^{-1}).$$

From the previous proposition, we know that both

$$\Phi(g) \circ \Phi(g^{-1}) \text{ and } \Phi(g^{-1}) \circ \Phi(g) \qquad \text{equal} \qquad \Phi(g^{-1} \cdot g) = \Phi(g \cdot g^{-1}) = \Phi(e).$$

So we only need to show that $\Phi(e) = \mathbf{1}_G$, or in words, that left multiplication by the identity does nothing. But this is obviously the case, in symbols:

$$\Phi(e)(x) = e \cdot x = x = \mathbf{1}_G(x).$$

■

Example B.5. Let's see how this theorem helps us realize the quaternion group of order 8, Q_8, as a subgroup of $\mathcal{S}_8$. Let's label the elements of Q_8 according to the following dictionary:

$\mathbf{1} = 1$	$-\mathbf{1} = 2$	$\mathbf{i} = 3$	$-\mathbf{i} = 4$
$\mathbf{j} = 5$	$-\mathbf{j} = 6$	$\mathbf{k} = 7$	$-\mathbf{k} = 8.$

[III] Be careful!! We have already shown that the function Φ is injective, but we are now trying to show that each $\Phi(g)$ is bijective! In fact the only groups for which Φ is bijective are the trivial groups and $\mathbb{Z}_2$!

Then, what does left multiplication by 1 do? It takes everybody to themselves. How about i? It behaves as follows:

$$1 \mapsto 3 \qquad 2 \mapsto 4 \qquad 3 \mapsto 2 \qquad 4 \mapsto 1$$
$$5 \mapsto 7 \qquad 6 \mapsto 8 \qquad 7 \mapsto 6 \qquad 8 \mapsto 5$$

(for instance, I got the $6 \mapsto 8$ statement by noting that 6 is $-\mathbf{j}$ and $\mathbf{i} \cdot (-\mathbf{j}) = -\mathbf{k}$ which is represented by the number 8). This means that

$$\Phi(\mathbf{i}) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & 2 & 1 & 7 & 8 & 6 & 5 \end{pmatrix} \quad \text{and} \quad \Phi(\mathbf{j}) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 6 & 8 & 7 & 2 & 1 & 3 & 4 \end{pmatrix}.$$

Or if you've read Chapter 7, which deals with cycle notation, these are the cycles

$$(1\,3\,2\,4)(5\,7\,6\,8) \quad \text{and} \quad (1\,5\,2\,6)(3\,8\,4\,7).$$

EXERCISE B.1. In this exercise I hope to further convince you that we have just made a copy of $\mathcal{Q}_8$ sitting inside of $\mathcal{S}_8$. Verify that just like the in $\mathcal{Q}_8$ we have that ($\Phi(\mathbf{i})$ and $\Phi(\mathbf{j})$ written above)

$$\Phi(\mathbf{i}) \circ \Phi(\mathbf{i}) = \Phi(\mathbf{j}) \circ \Phi(\mathbf{j}), \qquad |\Phi(\mathbf{i})| = 4, \qquad \Phi(\mathbf{j}) \circ \Phi(\mathbf{i}) \circ \Phi(\mathbf{j})^{-1} = \Phi(\mathbf{i})^{-1}.$$

SOLUTION

EXERCISE B.2. Write down the Cayley representation of $\mathbf{i}$ and χ_6 in $B\mathcal{D}_3$. (Be clear of what dictionary you are using!)

SOLUTION

C Reflections and Rotations

C.1 Orthogonal Transformation

Definition C.1. A matrix square matrix $\mathbf{A}$ is an *orthogonal transformation* if $\mathbf{A}^\top \cdot \mathbf{A}$ is the identity matrix. The set of n by n orthogonal transformations is denoted by $\mathbf{O}_n(\mathbb{R})$ This is called the *orthogonal group* of degree n.

Notice that by definition, a matrix is orthogonal iff it's transpose is its inverse. In particular, such matrices are always invertible. The following proposition shows that $\mathbf{O}_n(\mathbb{R})$ is closed under multiplication and inversion. (And hence, if you've read Chapter 6 you know that this means that $\mathbf{O}_n(\mathbb{R})$ is a group.)

Proposition C.2. *Suppose that* $\mathbf{A}$ *and* $\mathbf{B}$ *are elements of* $\mathbf{O}_n(\mathbb{R})$. *Then* $\mathbf{A} \cdot \mathbf{B}$ *and* $\mathbf{A}^{-1}$ *are as well.*[I]

Proof.

(1) Suppose that $\mathbf{A}^\top = \mathbf{A}^{-1}$ and $\mathbf{B}^\top = \mathbf{B}^{-1}$. We need to prove that $(\mathbf{A} \cdot \mathbf{B})^\top = \mathbf{B}^\top \cdot \mathbf{A}^\top$. Indeed

$$(\mathbf{A}\,\mathbf{B})^\top = \mathbf{B}^\top \mathbf{A}^\top = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

But by Socks and Shoes this is the inverse of $\mathbf{A}\,\mathbf{B}$. Hence this product is in $\mathbf{O}_n(\mathbb{R})$.

(2) Now we show that $\mathbf{A}^{-1} \in \mathbf{O}_n(\mathbb{R})$. Let's denote $\mathbf{A}^{-1}$ as $\mathbf{C}$. We need to prove that $\mathbf{C}$ is also in $\mathbf{O}_n(\mathbb{R})$. Then since $\mathbf{A}^\top \cdot \mathbf{A} = \mathbf{I}_n$, we have that $\mathbf{A}^\top = \mathbf{A}^{-1}$ and so with our notation

$$\mathbf{C} = \mathbf{A}^\top.$$

Therefore, $\mathbf{C}^\top = (\mathbf{A}^\top)^\top = \mathbf{A}$ and we have

$$\mathbf{C}^\top \cdot \mathbf{C} = \mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I}_n$$

as desired.

■

[I] Thus, if you've already read the chapter on subgroups, Chapter 6, we have proven that $\mathbf{O}_n(\mathbb{R})$ is a subgroup [hence a group] of $\mathbf{GL}_n(R)$.

EXERCISE C.1. Prove or give a counterexample: Is $\mathbf{O}_n(\mathbb{R})$ closed under addition?

SOLUTION

PROPOSITION C.3. *Every element of* $\mathbf{O}_n(\mathbb{R})$ *has determinant* $+1$ *or* -1.

PROOF. For any square matrix $\mathbf{A}$ we have that But by Theorem 3.205 and Theorem 3.212**(5)** we know that this implies that

$$\det \mathbf{A}^\top \mathbf{A} = \det \mathbf{A}^\top \det \mathbf{A} = \det \mathbf{A} \cdot \det \mathbf{A} = (\det \mathbf{A})^2 .$$

But if $\mathbf{A}$ is orthogonal we would have, by definition, $\mathbf{A}^\top \mathbf{A} = \mathbf{I}_n$, which has determinant 1. Therefore,

$$(\det \mathbf{A})^2 = 1 \Rightarrow \det \mathbf{A} = \pm 1.$$

as desired. ■

In a sense half have determinant 1 and half have determinant -1.

DEFINITION C.4. The *special orthogonal group* of degree n is the set of orthogonal transformations of determinant 1. That is

$$\mathbf{SO}_n(\mathbb{R}) \stackrel{\text{def}}{=} \{\mathbf{A} \in \mathbf{O}_n(\mathbb{R}) : \det \mathbf{A} = 1\}$$

Other ways of writing this include

$$\{\mathbf{A} \in \mathbf{M}_n(R) : \mathbf{A}^\top \cdot \mathbf{A} = \mathbf{I}_n \text{ and } \det \mathbf{A} = 1\} \quad \text{and} \quad \mathbf{SL}_n(R) \cap \mathbf{O}_n(\mathbb{R}) .$$

If you've already read about the index of subgroup in a group, then you should note that $\mathbf{SO}_n(\mathbb{R})$ has index 2 in $\varnothing[n]$.

EXAMPLE C.5. I claim that the matrix

$$\frac{1}{2}\begin{pmatrix} 1 & \phi^{-1} & -\phi \\ \phi^{-1} & \phi & 1 \\ \phi & -1 & \phi^{-1} \end{pmatrix}$$

is in $\mathbf{O}_3(\mathbb{R})$ where ϕ is the golden ratio, $2\cos \pi/5$. I do this by showing that 2 times this matrix has orthogonal colums, each of length 2. I also use the properties from Exercise 5.39.

$$\begin{aligned}
2\mathbf{c}_1 \bullet 2\mathbf{c}_1 &= \left\langle 1, \phi^{-1}, \phi \right\rangle \bullet \left\langle 1, \phi^{-1}, \phi \right\rangle = 1 + \phi^{-2} + \phi^2 = 1 + (2 - \phi) + (\phi + 1) = 4, \\
2\mathbf{c}_1 \bullet 2\mathbf{c}_2 &= \left\langle 1, \phi^{-1}, \phi \right\rangle \bullet \left\langle \phi^{-1}, \phi, -1 \right\rangle = \phi^{-1} + 1 - \phi = \phi - 1 + 1 - \phi = 0, \\
2\mathbf{c}_1 \bullet 2\mathbf{c}_3 &= \left\langle 1, \phi^{-1}, \phi \right\rangle \bullet \left\langle -\phi, 1, \phi^{-1} \right\rangle = -\phi + \phi^{-1} + 1 = 0, \\
2\mathbf{c}_2 \bullet 2\mathbf{c}_2 &= \left\langle \phi^{-1}, \phi, -1 \right\rangle \bullet \left\langle \phi^{-1}, \phi, -1 \right\rangle = \phi^{-2} + \phi^2 + 1 = 4, \\
2\mathbf{c}_2 \bullet 2\mathbf{c}_3 &= \left\langle \phi^{-1}, \phi, -1 \right\rangle \bullet \left\langle -\phi, 1, \phi^{-1} \right\rangle = -1 + \phi - \phi^{-1} = 0, \\
2\mathbf{c}_3 \bullet 2\mathbf{c}_3 &= \left\langle -\phi, 1, \phi^{-1} \right\rangle \bullet \left\langle -\phi, 1, \phi^{-1} \right\rangle = \phi^2 + 1 + \phi^{-2} = 4.
\end{aligned}$$

EXERCISE C.2. Prove that an orthogonal matrix is symmetric (i.e. $\mathbf{A}^\top = \mathbf{A}$) iff it is an involution.
Hint! This is more an exercise in definitions than anything else.

SOLUTION

EXERCISE C.3. Prove that $\mathbf{SO}_n(\mathbb{R})$ is also closed under multiplication and inversion. Show for any $\mathbf{A} \in \mathbf{O}_n(\mathbb{R})$ that $\mathbf{A}^2 \in \mathbf{SO}_n(\mathbb{R})$.[II]

SOLUTION

EXERCISE C.4. Show that

$$\mathbf{SO}_2(\mathbb{R}) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : \theta \in \mathbb{R} \right\}$$

(which are the set of all rotation matrices in $\mathbb{R}^2$, see (3).)
Hint! Hint, start with the equality

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

SOLUTION

THEOREM C.6. *Let* $\mathbf{A} \in \mathbf{M}_n(R)$. *Then* TFAE

(1) $\mathbf{A} \in \mathbf{O}_n(\mathbb{R})$,

(2) *for any two vectors* $\mathbf{v}$ *and* $\mathbf{w}$ *in* $\mathbb{R}^n$ *we have that*

$$\mathbf{A}(\mathbf{v}) \bullet \mathbf{A}(\mathbf{w}) = \mathbf{v} \bullet \mathbf{w},$$

(3) *The column vectors of the matrix* $\mathbf{A}$ *are orthonormal* (*Definition A.10 back on page 746*),

(4) The row vectors of the matrix $\mathbf{A}$ are orthonormal.

PROOF.

1 ⇒ 2 By definition, we are assuming that $\mathbf{A}^\top\mathbf{A} = \mathbf{I}_n$. Using this and 3.216 we have

$$\mathbf{A}(\mathbf{v}) \bullet \mathbf{A}(\mathbf{w}) = \left(\mathbf{A}^\top\mathbf{A}\mathbf{v}\right) \bullet \mathbf{w} = \mathbf{I}_n(\mathbf{v}) \bullet \mathbf{w} = \mathbf{v} \bullet \mathbf{w}$$

which is the statement of **(2)**.

2 ⇒ 3 Let $\mathbf{c}_1, \ldots, \mathbf{c}_n$ be the n column vectors of $\mathbf{A}$. Since the i^{th} column vector, $\mathbf{c}_i$ of any matrix $\mathbf{A}$ is simply the vector $\mathbf{A}(\mathbf{e}_i)$ we have that

$$\mathbf{c}_i \bullet \mathbf{c}_j = \mathbf{A}(\mathbf{e}_i) \bullet \mathbf{A}(\mathbf{e}_j) = \mathbf{e}_i \bullet \mathbf{e}_j = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise} \end{cases}$$

[II] Note, if you've read after Chapter 10 you know this is true simply because $\mathbf{SO}_n(\mathbb{R})$ has index 2 in $\mathbf{O}_n(\mathbb{R})$.

3 ⇔ 1 Let $\mathbf{c}_1, \dots, \mathbf{c}_n$ again denote the column vectors of $\mathbf{A}$. Then, as for any matrix,

$$\begin{aligned}(\mathbf{A}^\top\mathbf{A})(i,j) &= \sum_{j=1}^{n} A^\top_{i,j}A_{j,k}\\ &= A_{j,i}A_{j,k}\\ &= \sum_{j=1}^{n}(\mathbf{c}_i)_j(\mathbf{c}_k)_j\\ &= \mathbf{c}_i \bullet \mathbf{c}_j.\end{aligned}$$

Therefore, $(\mathbf{A}^\top\mathbf{A}) = \mathbf{I}_n$ iff the columns of $\mathbf{A}$ are orthonormal.

1 ⇔ 4 The statement $\mathbf{A}^\top\mathbf{A} = \mathbf{I}_n$ is the statement that $\mathbf{A}^{-1} = \mathbf{A}^\top$. Therefore $\mathbf{A}\mathbf{A}^\top = \mathbf{I}_n$ as well (since any matrix commutes with its inverse). Since we have already seem that 1. ⇔ 3. this implies that the columns of $\mathbf{A}^\top$ form an orthonormal basis. But this columns of $\mathbf{A}^\top$ are the rows of $\mathbf{A}$.

4 ⇒ 1 We have already seen that 4 is equivalent to the statement that $\mathbf{A}\mathbf{A}^\top = \mathbf{I}_n$. But for matrices this is equivalent to the statement that $\mathbf{A}^\top\mathbf{A} = \mathbf{I}_n$ completing the proof. ■

Since for any vector $\mathbf{v}$ we have that $\|\mathbf{v}\| = \sqrt{\mathbf{v}\bullet\mathbf{v}}$, every orthogonal transformation therefore fixes the length of any vector. as well as the distance between two points (vectors), since

$$\begin{gathered}\text{Distance between } \mathbf{v} \text{ and } \mathbf{w}\\ = d(\mathbf{v},\mathbf{w}) = \sqrt{(\mathbf{v}-\mathbf{w})\bullet(\mathbf{v}-\mathbf{w})} = d(\mathbf{A}(\mathbf{v}),\mathbf{A}(\mathbf{w}))\\ = \text{Distance between } \mathbf{A}(\mathbf{v}) \text{ and } \mathbf{A}(\mathbf{w}).\end{gathered}$$

C.2 Connection with Rigid Motions*

Recall that the distance between two points $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^n$ is measured by

$$\|\mathbf{v}-\mathbf{w}\| = \sqrt{(v_1-w_1)^2+\cdots+(v_n-w_n)^2}.$$

Indeed, this is nothing more than the Pythagorean theorem which can be extended to hold in any dimension.

Notation C.7. This distance between to points in $\mathbb{R}^n$ is denoted as

$$d(\mathbf{v},\mathbf{w}).$$

Remark C.8. In terms of lengths, we have that

$$d(\mathbf{v},\mathbf{w}) = \|\mathbf{v}-\mathbf{w}\|$$

and in terms of the dot product, we have that

$$d(\mathbf{v},\mathbf{w})^2 = (\mathbf{v}-\mathbf{w})\bullet(\mathbf{v}-\mathbf{w}).$$

If we were in analysis class, we'd spend more time talking about this distance function. Indeed, it satisfies the following three properties

(1) (Non-degeneracy) $\forall\, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n\ d(\mathbf{v}, \mathbf{w}) = 0$ iff $\mathbf{v} = \mathbf{w}$,

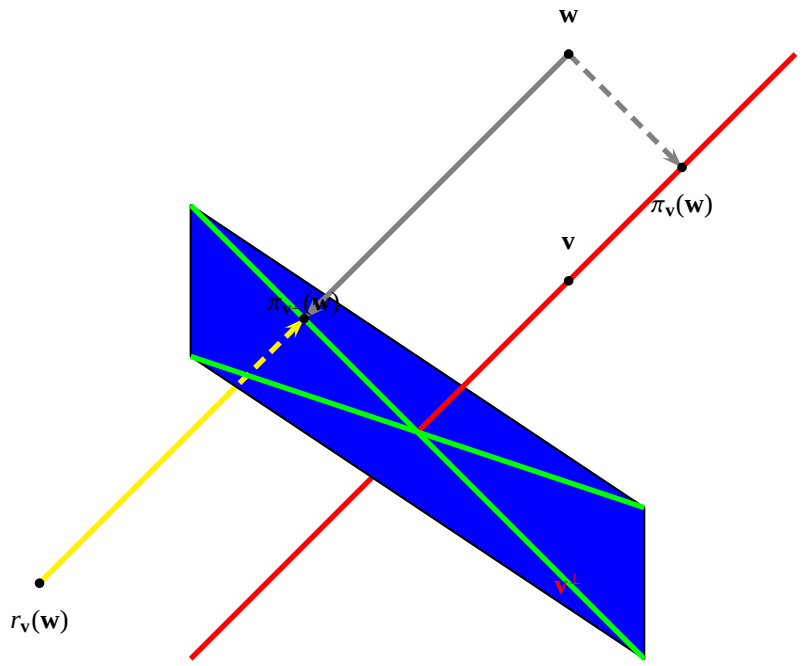
(2) (Symmetry) $\forall\, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n\ d(\mathbf{v}, \mathbf{w}) = d(\mathbf{w}, \mathbf{v})$;

(3) (Triangle Inequality) $\forall\, \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n\ d(\mathbf{v}, \mathbf{w}) \leq d(\mathbf{v}, \mathbf{u}) + d(\mathbf{u}, \mathbf{w})$.

The only property of the three which takes some (algebraic) proving is the third. The best logical path to it is to first prove the Cauchy-Schwartz inequality for the dot product, and then use this to prove this so-called triangle inequality. However, geometrically it should be pretty clear. In a sense, it says that the shortest distance between two points is a straight line.

C.3 REFLECTIONS

Let $\mathbf{w}, \mathbf{v} \in \mathbb{R}^3$ be two vectors (i.e. elements or points in $\mathbb{R}^3$) in 3 dimensional real space. In a Calculus 3 or Linear Algebra class one learns how to decompose $\mathbf{w}$ as the sum of two vectors, one perpendicular to $\mathbf{v}$ and the other parallel with it. We review this first.



The vector $\mathbf{w}$ can be written uniquely in the form

$$\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2$$

where

$$\mathbf{w}_1 \in \langle\, \mathbf{v} \,\rangle \quad \text{and} \quad \mathbf{w}_2 \in \mathbf{v}^{\perp \mathrm{I}}$$

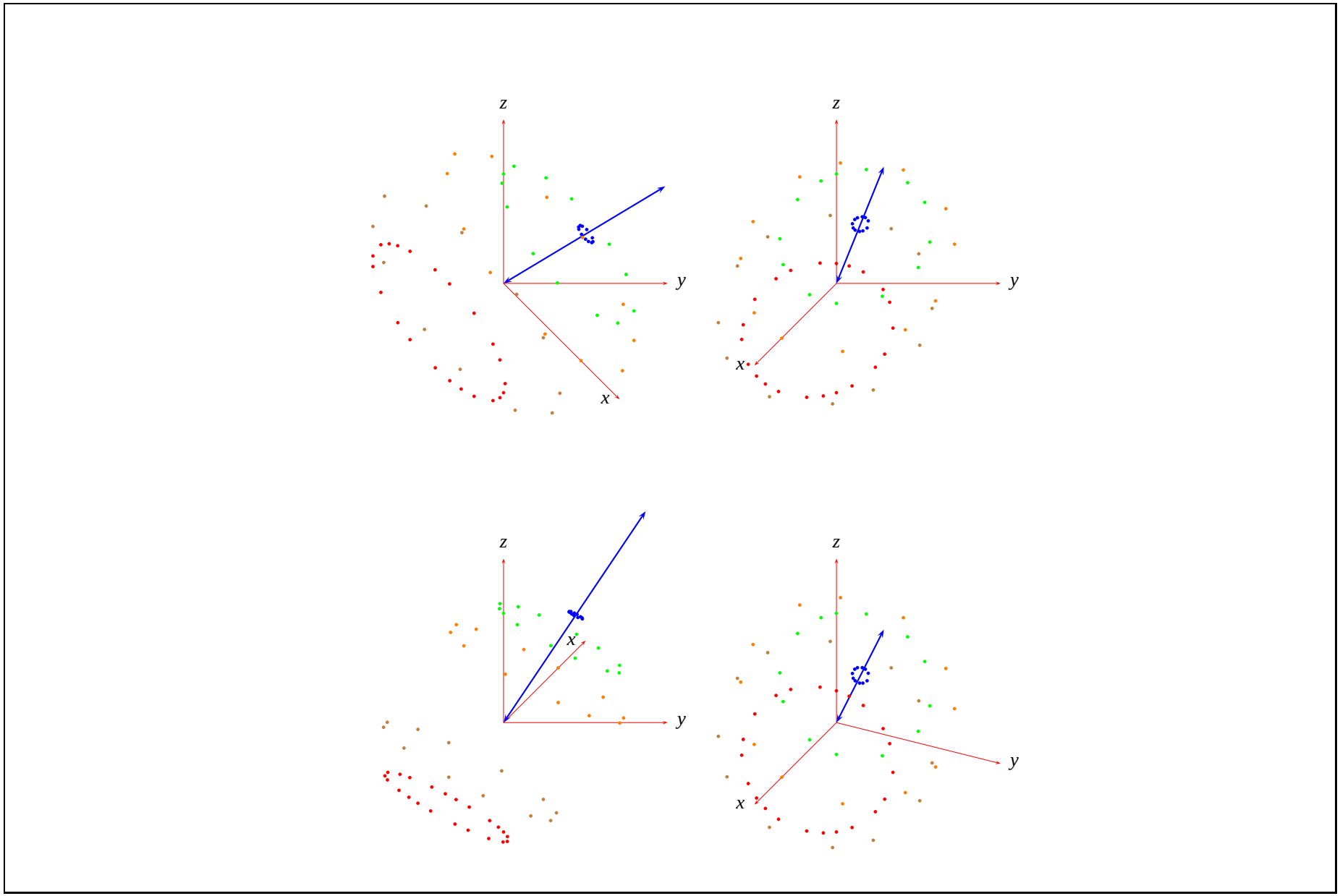


Table C.1: The rotation $\varrho_{\frac{2\pi}{3},\mathbf{j}} \circ \varrho_{\frac{2\pi}{3},\mathbf{i}}$ from four points of view.

This means we have well defined functions

$$\pi_{\mathbf{v}} : \mathbb{R}^n \mapsto \langle \mathbf{v} \rangle \qquad \pi_{\mathbf{v}^\perp} : \mathbb{R} \to \mathbb{R}^n$$
$$: \mathbf{w} \mapsto \mathbf{w}_1 \qquad : \mathbf{w} \mapsto \mathbf{w}_2.$$

[1] If you don't believe this, you can, in a sense, work this section backwards starting with the formulas! I assume you're more likely to believe this than the abstract forulas however.

Since $\pi_{\mathbf{v}}(\mathbf{w})$ is parallel to $\mathbf{v}$ there is some constant $c \in \mathbb{R}$ so that

$$\pi_{\mathbf{v}}(\mathbf{w}) = c\mathbf{v} \tag{1}$$

and since it is orthogonal to $\pi_{\mathbf{v}^\perp}(\mathbf{w})$ we have $\pi_{\mathbf{v}^\perp}(\mathbf{w}) \bullet \mathbf{v} = 0$. We use these facts to solve for c by dotting both sides of the equation

$$\mathbf{w} = \pi_{\mathbf{v}^\perp}(\mathbf{w}) + \pi_{\mathbf{v}}(\mathbf{w}) = \pi_{\mathbf{v}^\perp}(\mathbf{w}) + c\mathbf{v}$$

with $\mathbf{v}$. Then we end up with

$$\mathbf{v} \bullet \mathbf{w} = \mathbf{v} \bullet \pi_{\mathbf{v}^\perp}(\mathbf{w}) + c\mathbf{v} \bullet \mathbf{v} = c(\mathbf{v} \bullet \mathbf{v}).$$

Dividing both sides by $\mathbf{v} \bullet \mathbf{v}$ we see that

$$c = \frac{\mathbf{v} \bullet \mathbf{w}}{\mathbf{v} \bullet \mathbf{v}}.$$

Plugging in this c value into (1), we conclude the following:

PROPOSITION C.9. *For any vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ the projection of $\mathbf{v}$ onto $\mathbf{w}$ is given by*

$$\pi_{\mathbf{v}}(\mathbf{w}) = \frac{(\mathbf{v} \bullet \mathbf{w})}{(\mathbf{v} \bullet \mathbf{v})}\mathbf{v} \quad \textit{and} \quad \pi_{\mathbf{v}^\perp}(\mathbf{w}) = \mathbf{w} - \frac{(\mathbf{v} \bullet \mathbf{w})}{(\mathbf{v} \bullet \mathbf{v})}\mathbf{v}.$$

■

From the formulas, we easily see that $\pi_{\mathbf{v}}$ and $\pi_{\mathbf{v}^\perp}$ are linear. Indeed, the first is a composition of linear functions,

$$\pi_{\mathbf{v}} = \iota_{\frac{\mathbf{v}}{\|\mathbf{v}\|^2}} \circ (\mathbf{v} \bullet),$$

where

$$\begin{aligned} \mathbf{v} \bullet &: \mathbb{R}^n \mapsto \mathbb{R} \\ &: \mathbf{w} \mapsto \mathbf{v} \bullet \mathbf{w} \end{aligned} \qquad \begin{aligned} \iota_{\mathbf{u}} &: \mathbb{R} \to \mathbb{R}^n \\ &: c \mapsto c\mathbf{u} \end{aligned}$$

and the second is the difference of linear functions $\mathbf{1}_{\mathbb{R}^n} - \pi_{\mathbf{v}}$.[II]

EXERCISE C.5. For $\mathbf{v} \neq \mathbf{0}$ prove

(1) Using the above formula verify that if $\mathbf{v} \asymp \mathbf{v}'$, with neither $\mathbf{0}$, then $\pi_{\mathbf{v}} = \pi_{\mathbf{v}'}$

(2) Show that

$$\pi_{\mathbf{v}} \circ \pi_{\mathbf{v}} = \pi_{\mathbf{v}}.$$

(3) Show that

$$\pi_{\mathbf{v}}(\mathbf{w}) = \mathbf{0} \Leftrightarrow \mathbf{v} \perp \mathbf{w}.$$
$$\pi_{\mathbf{v}}(\mathbf{w}) = \mathbf{w} \Leftrightarrow \mathbf{v} \asymp \mathbf{w}.$$

SOLUTION

[II] Since

$$\mathbf{v} \bullet (\mathbf{w} - \frac{(\mathbf{v} \bullet \mathbf{w})}{(\mathbf{v} \bullet \mathbf{v})}\mathbf{v}) = \mathbf{v} \bullet \mathbf{w} - \frac{(\mathbf{v} \bullet \mathbf{w})}{(\mathbf{v} \bullet \mathbf{v})}\mathbf{v} \bullet \mathbf{v} = \mathbf{v} \bullet \mathbf{w} - \mathbf{v} \bullet \mathbf{w} = 0$$

we indeed see the claim that we started out with, i.e. that every vector can be uniquely written as $\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2$ was indeed true after all! Alas, we used this fact to derive the formulas in the first place.

If we factor $\mathbf{w}$ as $\mathbf{w} = \pi_{\mathbf{v}^\perp}(\mathbf{w}) + \pi_{\mathbf{v}}(\mathbf{w})$ the reflection of $\mathbf{w}$ leaves the component of $\mathbf{w}$ along the perpendicular direction of $\mathbf{v}$ and swaps the parallel direction. Thus

$$r_{\mathbf{v}}(\mathbf{w}) = \pi_{\mathbf{v}^\perp}(\mathbf{w}) - \pi_{\mathbf{v}}(\mathbf{w}) = \mathbf{v} - 2\pi_{\mathbf{v}}(\mathbf{w}).$$

Plugging in for $2\pi_{\mathbf{v}}(\mathbf{w})$ we arrive at the following formula:

THEOREM C.10. *For a vector* $\mathbf{v}$ *we therefore have the following formula for a reflection:*

$$r_{\mathbf{v}}(\mathbf{w}) = \mathbf{w} - 2\frac{(\mathbf{v} \bullet \mathbf{w})}{(\mathbf{v} \bullet \mathbf{v})}\mathbf{v}.$$

■

In particular, if we replace $\mathbf{v}$ with its normalization $\widehat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$, so $\widehat{\mathbf{v}} \bullet \widehat{\mathbf{v}} = 1$, we end up with

$$r_{\mathbf{v}}(\mathbf{w}) = r_{\widehat{\mathbf{v}}}(\mathbf{w}) = \mathbf{w} - 2(\widehat{\mathbf{v}} \bullet \mathbf{w})\widehat{\mathbf{v}}. \tag{2}$$

THEOREM C.11.

(1) *Reflection is a linear transformation.*

(2) $r_{\mathbf{v}}(\mathbf{v}) = -\mathbf{v}$.

(3) $r_{\mathbf{v}}$ *fixes only the vectors orthogonal to* $\mathbf{v}$. *In symbols*

$$r_{\mathbf{v}}(\mathbf{w}) = \mathbf{w} \Leftrightarrow \mathbf{w} \perp \mathbf{v}.$$

(4) *Reflection is an involution, that is* $r_{\mathbf{v}} \circ r_{\mathbf{v}} = \mathbf{1}$. [III]

(5) *Reflection is orthogonal, that is for any* $\mathbf{w}_1$ *and* $\mathbf{w}_2$ *we have*

$$r_{\mathbf{v}}(\mathbf{w}_1) \bullet r_{\mathbf{v}}(\mathbf{w}_2) = \mathbf{w}_1 \bullet \mathbf{w}_2.$$

(6) *For non-parallel vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *the only fixed points of*

$$r_{\mathbf{u}} \circ r_{\mathbf{v}}$$

are points orthogonal to both $\mathbf{u}$ *and* $\mathbf{v}$.

(7) *The determinant of any reflection is* (-1).

PROOF.

(1) Since

$$r_{\mathbf{v}}(\mathbf{w}) = \mathbf{w} - 2\pi_{\mathbf{v}}(\mathbf{w}) \Rightarrow r_{\mathbf{v}} = \mathbf{1} - 2\pi_{\mathbf{v}}$$

a reflection is a linear combination of linear transformations, and is hence a linear transformation.

(2) $\pi_{\mathbf{v}}(\mathbf{v}) = \mathbf{v}$ and $\pi_{\mathbf{v}^\perp}(\mathbf{v}) = \mathbf{0}$. Hence $\mathbf{v} = \mathbf{0} + \pi_{\mathbf{v}}(\mathbf{v})$ and therefore the result.

(3) If $r_{\mathbf{v}}(\mathbf{w}) = \mathbf{w}$, then

$$\begin{aligned}\pi_{\mathbf{v}^\perp}(\mathbf{w}) + \pi_{\mathbf{v}}(\mathbf{w}) = \mathbf{w} = \mathcal{R}_{\mathbf{v}}(\mathbf{w}) = \pi_{\mathbf{v}^\perp}(\mathbf{w}) - \pi_{\mathbf{v}}(\mathbf{w}) \\ \Leftrightarrow 2\pi_{\mathbf{v}}(\mathbf{w}) = \mathbf{0} \\ \Leftrightarrow \pi_{\mathbf{v}}(\mathbf{w}) = \mathbf{0} \Leftrightarrow \mathbf{w} \in \mathbf{v}^\perp.\end{aligned}$$

[III] No, I didn't put $\circ$ instead of $\bullet$ instead! I really mean any reflection composed with itself is the identity.

(4) Clearly this ought to hold if we have the correct definition (If you reflection something twice you've down nothing ... the man in the mirror sees **you** in the mirror!) We have

$$\begin{aligned} r_{\mathbf{v}}(r_{\mathbf{v}}(\mathbf{w})) &= r_{\mathbf{v}}(\pi_{\mathbf{v}^\perp}(\mathbf{w}) - \pi_{\mathbf{v}}(\mathbf{w})) \\ &= r_{\mathbf{v}}(\pi_{\mathbf{v}^\perp}(\mathbf{w})) - r_{\mathbf{v}}(\pi_{\mathbf{v}}(\mathbf{w})) \\ &= \pi_{\mathbf{v}^\perp}(\mathbf{w}) + \pi_{\mathbf{v}}(\mathbf{w}) \\ &= \mathbf{w}. \end{aligned}$$

(5) Let $\mathbf{u}$, $\mathbf{w}$ be arbitrary. Then

$$\begin{aligned} r_{\mathbf{v}}(\mathbf{u}) &= \pi_{\mathbf{v}^\perp}(\mathbf{u}) - \pi_{\mathbf{v}}(\mathbf{u}) \quad \text{and} \quad r_{\mathbf{v}}(\mathbf{w}) \\ &= \pi_{\mathbf{v}^\perp}(\mathbf{w}) - \pi_{\mathbf{v}}(\mathbf{w}) \end{aligned}$$

$$\begin{aligned} \Rightarrow r_{\mathbf{v}}(\mathbf{u}) \bullet r_{\mathbf{v}}(\mathbf{w}) &= (\pi_{\mathbf{v}^\perp}(\mathbf{u}) - \pi_{\mathbf{v}}(\mathbf{u})) \bullet (\pi_{\mathbf{v}^\perp}(\mathbf{w}) - \pi_{\mathbf{v}}(\mathbf{w})) \\ &= \pi_{\mathbf{v}^\perp}(\mathbf{u}) \bullet \pi_{\mathbf{v}^\perp}(\mathbf{w}) + \pi_{\mathbf{v}}(\mathbf{u}) \bullet \pi_{\mathbf{v}}(\mathbf{w}) \\ &= \pi_{\mathbf{v}^\perp}(\mathbf{u}) \bullet \pi_{\mathbf{v}^\perp}(\mathbf{w}) + \pi_{\mathbf{v}}(\mathbf{u}) \bullet \pi_{\mathbf{v}}(\mathbf{w}) \\ &= (\pi_{\mathbf{v}^\perp}(\mathbf{u}) + \pi_{\mathbf{v}}(\mathbf{u})) \bullet (\pi_{\mathbf{v}^\perp}(\mathbf{w}) + \pi_{\mathbf{v}}(\mathbf{w})) \\ &= \mathbf{u} \bullet \mathbf{w}. \end{aligned}$$

(6) Select an orthogonal basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ with $\mathbf{b}_1$ proportional to $\mathbf{v}$. (In $\mathbb{R}^3$ we can see that this can be done by first selecting a vector which is neither parallel nor perpendicular to $\mathbf{v}$, projecting it to $\mathbf{v}$ to get $\mathbf{b}_1$, then projecting the original vector to $\mathbf{v}^\perp$ to get $\mathbf{b}_2$. Lastly, take the cross product of $\mathbf{b}_1$ and $\mathbf{b}_2$ to get $\mathbf{b}_3$. Since we started with a vector neither perpendicular nor parallel to $\mathbf{v}$ none of these vectors are zero.) It's clearly that the determinant of $(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ is nonzero (in fact it's an easy exercise to show that, up to sign, this determinant is the product of the lengths of the $\mathbf{b}_i$'s.) Now, for our reflection $r_{\mathbf{v}}$, we know that $r_{\mathbf{v}}(\mathbf{b}_1) = -\mathbf{b}_1$ since $\mathbf{b}_1$ is proportional to $\mathbf{v}$, and for all $i > 1$ we have that $r_{\mathbf{v}}(\mathbf{b}_i) = \mathbf{b}_i$, since such vectors are parallel to $\mathbf{v}$. Therefore

$$\begin{aligned} \det r_{\mathbf{v}} \cdot \det(\mathbf{b}_1, \ldots, \mathbf{b}_n) &= \det(r_{\mathbf{v}}(\mathbf{b}_1), \ldots, r_{\mathbf{v}}(\mathbf{b}_n)) \\ &= \det(-\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n) \\ &= -\det(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n). \end{aligned}$$

Canceling out the non-zero $\det(\mathbf{b}_1, \ldots, \mathbf{b}_n)$ proves that $\det r_{\mathbf{v}} = -1$ as desired. ■

The last part of this theorem can be proven in one line if we assumed some basic linear algebra. Simply note that the eigenvalues of a reflections are all equal to 1 except for a single eigenvalue which is equal to -1. However, since we haven't gone into eigenvalues or eigenvectors, we gave the more elementary proof above.

For two vectors $\mathbf{u}$ and $\mathbf{v}$ at least one of $\mathbf{u} \bullet \mathbf{v}$ or $(-\mathbf{u}) \bullet \mathbf{v}$ must be positive and be equal to $|\mathbf{u} \bullet \mathbf{v}|$. For any two vectors $\mathbf{u}$ and $\mathbf{v}$ we define the reduced angle between $\mathbf{u}$ and $\mathbf{v}$ as

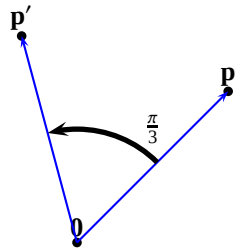
$$\cos^{-1}(|\mathbf{u} \bullet \mathbf{v}|).$$

Evidentially this measures the angle between the two hyperplanes $\mathbf{u}^\perp$ and $\mathbf{v}^\perp$.

C.4 Rotations and Reflections in the Plane

DEFINITION C.12. A *rotation* in the plane (about the origin) by an angle θ is a transformation ρ which

- $\varrho_\theta(\mathbf{0}) = \mathbf{0}$,
- If $\mathbf{p} \neq \mathbf{0}$, then the orientated angle between $\mathbf{p}$ and $\varrho_\theta(\mathbf{p})$ is θ.



Notice that, for rotations in the plane, the rotation angle is additive. That is

$$\varrho_\theta \circ \varrho_\psi = \varrho_{\theta+\psi}.$$

EXAMPLE C.13. Let $\mathbf{v} = (2, 2)$ in the plane. The polar angle of this point is $\pi/4$ and so the polar angle of the new point is $\pi/4 + \pi/3$. Using double angle formulas we therefore see that

$$\varrho_{\pi/3}(2, 2) = (1 - \sqrt{3}, 1 + \sqrt{3}).$$

There are a couple of nice general formulas for this rotation.

(1) (Complex) In Section 3.2 of Chapter 3 (the section on complex numbers) we showed that you can realize rotation by an angle θ by multiplying by $e^{i\theta}$. In this sense, if we write a point (a, b) in the plane $\mathbb{R}^2$ as the complex number $a + bi$, then rotating this point by an angle θ is the same as multiplying the point $a + bi$ by the complex number

$$e^{i\theta} = \cos\theta + \sin\theta\, i.$$

For example, if $\theta = \pi/3$ as was the case with Example C.13 that we just did, then

$$e^{i\pi/3} = \cos\frac{\pi}{3} + \sin\frac{\pi}{3}\, i = \frac{1}{2} + \frac{\sqrt{3}}{2}\, i.$$

Thus, we can recover that example by performing the multiplication

$$\begin{aligned}\left(\frac{1}{2} + \frac{\sqrt{3}}{2}\, i\right) \cdot \left(2 + 2\, i\right) &= \left(1 + \sqrt{3}\, i\right) \cdot \left(1 + \ i\right) \\ &= \left(1 - \sqrt{3}\right) + \left(1 + \sqrt{3}\right)\, i.\end{aligned}$$

This again shows that $\varrho_{\pi/3}(2, 2) = (1 - \sqrt{3}, 1 + \sqrt{3})$, but this time we didn't need to use a double angle formula.

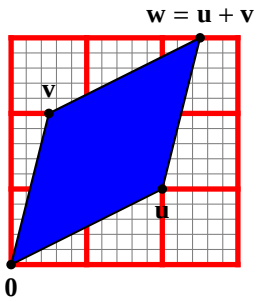
(2) There are a number of ways of seeing that a rotation is a linear transformation.

- You can express it as the composition of two reflections and use the fact that the composition of two linear maps is linear.
- You can use the formula in terms of complex numbers that we just came up with.
- However, the most intuitive way is to translate the conditions that a map be linear into geometrical language. Let's translate the two conditions a map F needs to be linear (in the special case of the plane, $\mathbb{R}^2$)

$$\forall \mathbf{v}, \mathbf{w} \in \mathbb{R}^2, \qquad F(\mathbf{v}+\mathbf{w}) = F(\mathbf{v}) + F(\mathbf{w}) \tag{L1}$$

$$\forall \mathbf{v} \in \mathbb{R}^2, c \in \mathbb{R}, \qquad F(c\,\mathbf{v}) = c\,F(\mathbf{v}). \tag{L2}$$

into geometric language. Recall, that three vectors $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$ in the plane satisfy $\mathbf{u}+\mathbf{v}=\mathbf{w}$ iff $\mathbf{w}$ lives on the fourth vertex of a parallelogram which contains the origin, $\mathbf{u}$ and $\mathbf{v}$.



Notice that a rotation clearly preserves this configuration. Therefore (two dimensional) rotations are additive.

To see that a rotation preserves scalar multiplication is along the same lines. For ease, we assume that c is a positive real number. The case that c is a negative real number is along the same lines (and actually can be proven directly from the additive property already proven.) Notice that the geometric significance of $c\mathbf{v}$ is that it is a point on the line containing the origin and $\mathbf{v}$ but which is a distance scaled by c. If we think of $\mathbf{v}$ in the typical vector as an arrow way, then $c\mathbf{v}$ is the arrow which is $\mathbf{v}$ but scaled accordingly. If we apply this geometric insight, it is clear that if we apply a rotation ρ (which preserves the origin) to $\mathbf{v}$, it will take $\rho(c\mathbf{v})$ to $c\rho(\mathbf{v})$.

- Looking ahead, we can apply Theorem C.27 which says that any rigid motion which preserves the origin is linear.

Anyway you choose to look at it, we see that

THEOREM C.14. *Rotations are linear transformations.* ■

In particular, they have a matrix representation which is determined uniquely by following what it does to the two standard basis vectors:

$$\mathbf{i} = (1,0) \qquad \mathbf{j} = (0,1).$$

Now, for a rotation by an angle θ it's easy to see what happens to these two points by staring at the unit circle for a moment, see Figure C.1.

Since

$$\cos\left(\theta + \frac{\pi}{2}\right) = \cos\theta \cdot \cos\frac{\pi}{2}^{0} - \sin\theta \cdot \sin\frac{\pi}{2}^{1} = -\sin\theta$$

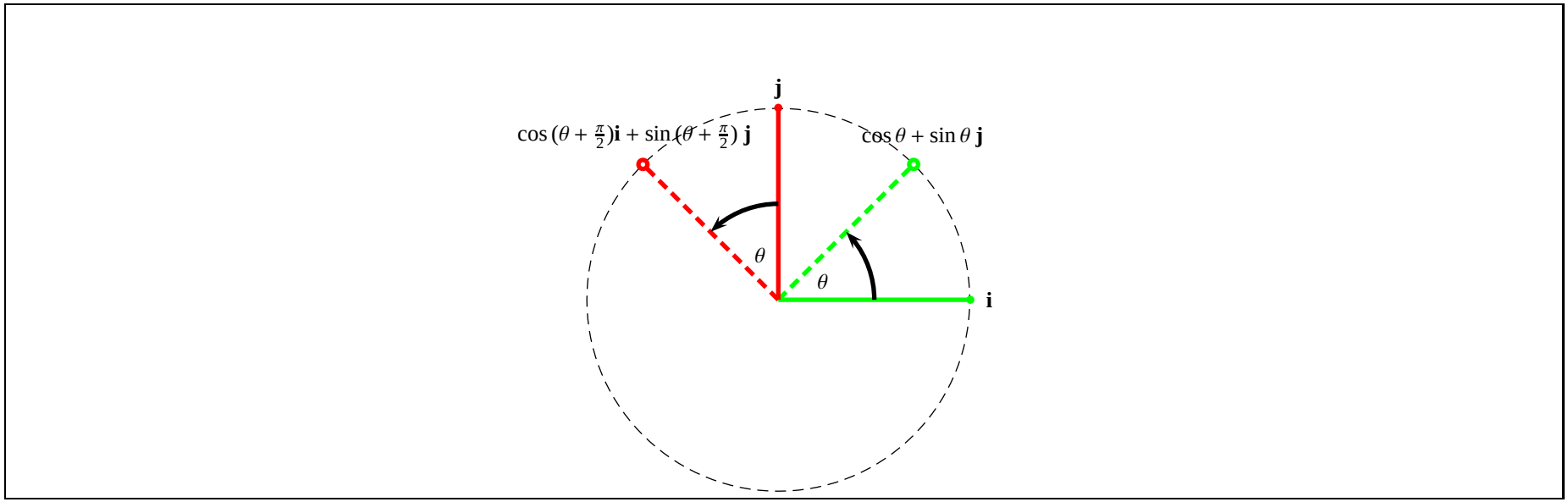


Figure C.1: A rotation in the plane.

and

$$\sin\left(\theta+\frac{\pi}{2}\right)=\cos\theta\cdot\sin\frac{\pi}{2}-\sin\theta\cdot\cos\frac{\pi}{2}=\cos\theta,$$

we can write ρ_θ as a matrix

$$\rho_\theta=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}. \tag{3}$$

C.5 Rotations in $\mathbb{R}^3$

Now we move up one dimension from the plane, $\mathbb{R}^2$, to the space of everyday life, $\mathbb{R}^3$. Note that if the fixed points of a rigid motion φ are a line, $\mathcal{L}$, and so in particular φ does not fix any point of $\mathcal{L}^\perp$, it must rotate this plane.

Definition C.15. A *rotation* in $\mathbb{R}^n$ is an orthogonal transformation φ which does not fix two non-proportional vectors **v** and **w** but fixes every point perpendicular to both **v** and **w**.

How do we denote a rotation in three dimensions? Clearly we can no longer simply give an angle and be done with the matter. Rotations in three dimensions depend on, not only an angle, but also a fixed axis of rotation. Unfortunately, it is also not quite enough to give an angle and a line since the rotation also depends on an orientation of the angle with respect to the line. For instance, if you turn a screw by angle in a "lefty" way, the screw will always move "towards the wall" and away from you. Now imagine a double sided screw hanging out two sides of a wall. A friend of yours standing on the other side of the wall also turns the screw by the same angle also in "lefty" direction. This also moves the screw towards the wall and away from him, but this is towards you, so the same angle and axis have the complete opposite effect.

As such, we need denote a rotation an angle, a line, **and** a choice of a meaning of what side of the line is "positive" and what side of the line is "negative." The best way to do this is by simply picking a vector on the line. Of course since any non-zero vector $\mathbf{v}$ on the line $\mathcal{L}$ determines the line, via

$$\mathcal{L} = \langle\, \mathbf{v} \,\rangle \stackrel{\text{def}}{=} \{c\mathbf{v} : c \in \mathbb{R}\}$$

a rotation is determined (but not uniquely) by an angle and a vector. As such, we use the notation

$$\varrho_{\theta,\mathbf{v}}$$

for a rotation by an angle θ along the line $\langle\, \mathbf{v} \,\rangle$ with a choice of the line so $\mathbf{v}$ lies on the positive half. This means that $\varrho_{\mathbf{v},\theta}$ "when looking down on $\mathbf{v}$, $\varrho_{\mathbf{v},\theta}$ is a positive (counter-clockwise) rotation by an angle θ." Another way to visualize is by taking your pointer finger in the direction of some initial $\mathbf{w}$ and curling your fingers and angle θ towards $\varrho_{\mathbf{v},\theta}(\mathbf{w})$ forces your thumb to point up in the direction of $\mathbf{v}$.

Notice that with our convention, a rotation in the plane ϱ_θ becomes the rotation $\varrho_{\theta,\mathbf{k}}$ when thought of as a rotation in 3-space.

As already pointed out, if we change the half of the line that we want to be positive, we end up taking the inverse of the rotation. (Some jerk ⟨ok, he's not your friend anymore! ⟩ standing on the other side of a wall turning counter-clockwise by an angle θ standing on the opposite side of the wall from you completely undoes what you did by doing a counter-clockwise turn ⟨like that fool⟩ by an angle θ ⟨also just like that fool⟩.) Symbolically, this means that

$$\varrho_{\theta,\mathbf{w}} = \varrho^{-1}_{\theta,\mathbf{v}}.$$

whenever $\mathbf{v}$ and $\mathbf{w}$ defer by a negative real number (meaning whenever there exists a negative number c so that $c\mathbf{v} = \mathbf{w}$.) In particular, since -1 is a negative number, we have

$$\varrho_{\theta,-\mathbf{v}} = \varrho^{-1}_{\theta,\mathbf{v}}.$$

Of course we also have that

$$\varrho_{\theta,\mathbf{v}} = \varrho_{\theta,\mathbf{w}}$$

whenever ever $\mathbf{v}$ and $\mathbf{w}$ differ by a positive real number.

Typically, since we have for any non-zero vector $\mathbf{u}$, $\|\,\mathbf{u}\,\|$ is a positive number, so that

$$\mathbf{u} \quad \text{and} \quad \widehat{\mathbf{u}} = \frac{\mathbf{u}}{\|\,\mathbf{u}\,\|}$$

differ by a positive number, we may as well (and we typically do) assume that we are dealing with a normalized vector, i.e. a vector of length 1, i.e. a vector in $\mathbb{S}^2$.

It's also worth noting that for a fixed vector, $\mathbf{v}$ we have and

$$\varrho_{\theta+\psi,\mathbf{v}} = \varrho_{\theta,\mathbf{v}} \circ \varrho_{\psi,\mathbf{v}} \text{ and } \varrho_{2\pi,\mathbf{v}} = \mathbf{1}$$

which are just the basic properties of rotations in the plane.

Let's get a more mathematical meaning of what we mean by "looking down on." This means that we have that for $\mathbf{v} \in \mathbb{R}^3$ and $0 \le \theta \le \pi$, we must have

$$\det\left(\mathbf{v}, \varrho_{\widehat{\mathbf{u}},\theta}(\mathbf{v}), \widehat{\mathbf{u}}\right) \ge 0. \tag{4}$$

See Figure C.3 back on page 771.

Actually, for every vector $\mathbf{v} \perp \widehat{\mathbf{u}}$ we have

$$\varrho_{\theta,\widehat{\mathbf{u}}}(\mathbf{v}) = \cos\theta\,\mathbf{v} + \sin\theta\,\widehat{\mathbf{u}} \times \mathbf{v}.^{\text{I}} \tag{5}$$

This can be seen from Figure C.2.

A rotation is therefore any isometry which fixes a line and rotates the plane perpendicular to the line.

[I] It's worth pointing out that

$$\mathbf{v} \times (\widehat{\mathbf{u}} \times \mathbf{v}) = (\mathbf{v} \bullet \mathbf{v}) \cdot \widehat{\mathbf{u}} - (\mathbf{v} \bullet \widehat{\mathbf{u}})\mathbf{v} = (\mathbf{v} \bullet \mathbf{v}) \cdot \widehat{\mathbf{u}}$$

and that $\left\|\widehat{\mathbf{u}} \times \mathbf{v}\right\| = \|\,\mathbf{v}\,\|$.

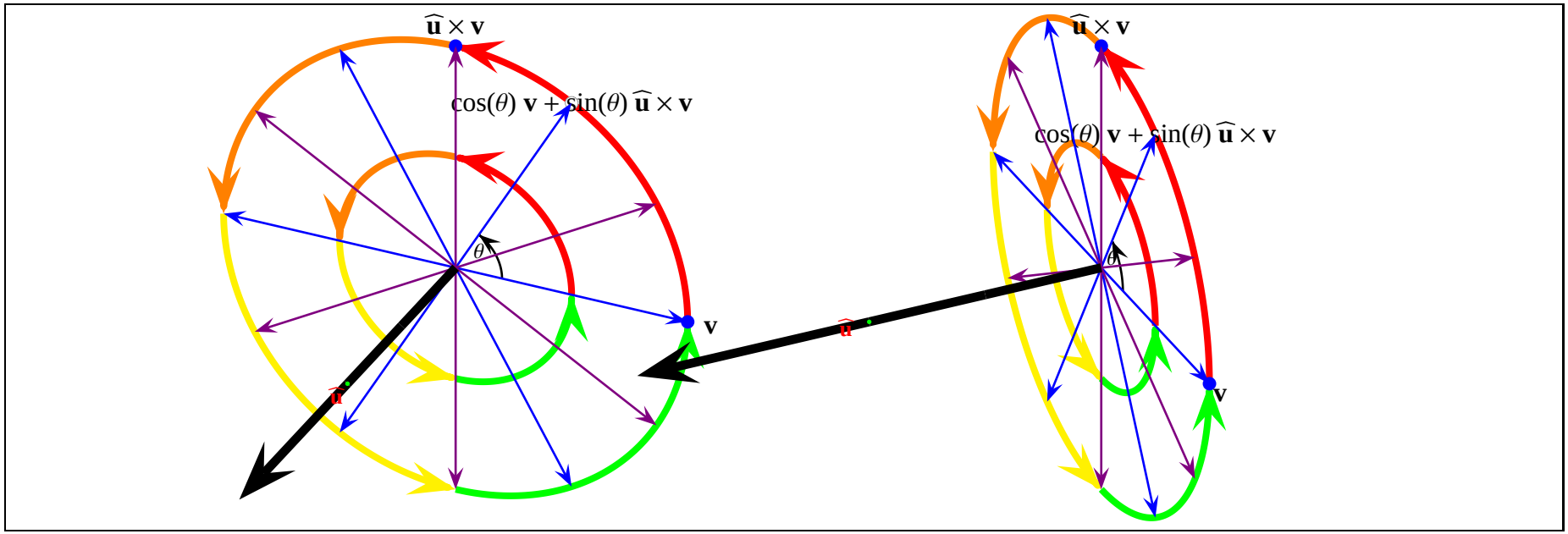


Figure C.2: Note this is a 3-dimensional version of Figure C.1. Simply replace **i** with arbitrary **v** and replace **j**–which is right-handed perpendicular to **i** relative to **k** – with $\mathbf{u} \times \mathbf{v}$–which is right-handed perpendicular to **v** relative to **u**.

EXERCISE C.6. Prove for all $\mathbf{v} \perp \widehat{\mathbf{u}}$

(1) $\det(\mathbf{v}, \varrho_{\theta,\widehat{\mathbf{u}}}(\mathbf{v}), \widehat{\mathbf{u}}) = \sin\theta \cdot \|\mathbf{v}\|^2$.

(2) $\mathbf{v} \times \varrho_{\theta,\widehat{\mathbf{u}}}\mathbf{v} = \sin\theta \|\mathbf{v}\|^2\, \mathbf{u}$.

SOLUTION

As $\widehat{\mathbf{u}} \times \mathbf{v} = \widehat{\mathbf{u}} \times \pi_{\widehat{\mathbf{u}}^\perp}(\mathbf{v})$, this means for an arbitrary $\mathbf{v} \in \mathbb{R}^3$ we have

$$\begin{aligned}\varrho_{\widehat{\mathbf{u}},\theta}(\mathbf{v}) &= \pi_{\widehat{\mathbf{u}}}(\mathbf{v}) + \cos\theta \cdot \pi_{\widehat{\mathbf{u}}^\perp}(\mathbf{v}) + \sin\theta \cdot \widehat{\mathbf{u}} \times \mathbf{v} \\ &= (\widehat{\mathbf{u}} \bullet \mathbf{v})\,\widehat{\mathbf{u}} + \cos\theta \cdot (\mathbf{v} - (\widehat{\mathbf{u}} \bullet \mathbf{v})\,\widehat{\mathbf{u}}) + \sin\theta \cdot \widehat{\mathbf{u}} \times \mathbf{v}.\end{aligned}$$

THEOREM C.16. *Let* **v** *and* **w** *be 2 non-parallel 3-vectors (hence both must be non-zero). Then,*

$$r_{\mathbf{w}} \circ r_{\mathbf{v}} = \varrho_{2\theta,\mathbf{v}\times\mathbf{w}}$$

where $\theta = \theta(\mathbf{v}, \mathbf{w})$ *is the angle (between* 0 *and* $\frac{\pi}{2}$*) between the lines* $\langle \mathbf{v} \rangle$ *and* $\langle \mathbf{w} \rangle$. [II]

PROOF.

[II] In other words, we have $\|\mathbf{v} \times \mathbf{w}\| = \sin\theta$.

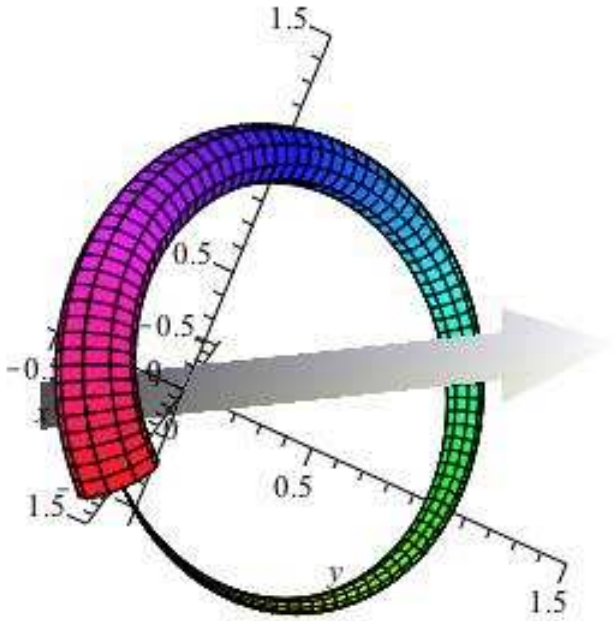


Figure C.3: A 3d Rotation This depicts the effect of $\varrho_{\theta,\mathbf{v}}$ on $\mathbf{i}$, where $\frac{\mathbf{i}+\mathbf{j}+\mathbf{k}}{\sqrt{3}}$, for increasing τ. I indicate the effect of larger τ by increasing the size of the tube. The vector $\frac{\mathbf{i}+\mathbf{j}+\mathbf{k}}{\sqrt{3}}$ is indicated.

(1) Define ρ as $r_{\mathbf{w}} \circ r_{\mathbf{v}}$. As the composition of orthogonal transformations, ρ is an orthogonal transformation.

(2) For every

$$\mathbf{u} \in \mathbf{v}^{\perp} \cap \mathbf{w}^{\perp}$$

both $r_{\mathbf{v}}$ and $r_{\mathbf{w}}$ fix $\mathbf{u}$, and hence ρ fixes $\mathbf{u}$. Conversely, if ρ fixes a vector $\mathbf{u}$, then using the assumption that $\mathbf{v}$ and $\mathbf{w}$ are not proportional:

$$\begin{aligned}(r_{\mathbf{w}} \circ r_{\mathbf{v}})\,\mathbf{u} = \mathbf{u} &\Rightarrow r_{\mathbf{v}}(\mathbf{u}) = r_{\mathbf{w}}(\mathbf{u}) \\ &\Rightarrow \mathbf{u} - 2\frac{\mathbf{u}\bullet\mathbf{v}}{\mathbf{v}\bullet\mathbf{v}}\mathbf{v} = \mathbf{u} - 2\frac{\mathbf{u}\bullet\mathbf{w}}{\mathbf{w}\bullet\mathbf{w}}\mathbf{w} \\ &\Rightarrow \frac{\mathbf{u}\bullet\mathbf{v}}{\mathbf{v}\bullet\mathbf{v}}\mathbf{v} = \frac{\mathbf{u}\bullet\mathbf{w}}{\mathbf{w}\bullet\mathbf{w}}\mathbf{w} \\ &\Rightarrow \frac{\mathbf{u}\bullet\mathbf{v}}{\mathbf{v}\bullet\mathbf{v}} = \frac{\mathbf{u}\bullet\mathbf{w}}{\mathbf{w}\bullet\mathbf{w}} = 0 \\ &\Rightarrow \mathbf{u}\bullet\mathbf{v} = 0 \quad \text{and} \quad \mathbf{u}\bullet\mathbf{w} = 0.\end{aligned}$$

Therefore $\mathbf{u}$ is perpendicular to both $\mathbf{v}$ and $\mathbf{w}$.

(3) So far we've proven that ρ is a rotation. Therefore, for any $\mathbf{u} \in \mathbf{v}^{\perp} \cap \mathbf{w}^{\perp}$, $r_{\mathbf{w}} \circ r_{\mathbf{v}} = \varrho_{\epsilon(\mathbf{u})\cdot\alpha\,,\mathbf{u}}$ for some angle α between 0 and π and some $\epsilon(\mathbf{u}) = \pm 1$ depending on $\mathbf{u}$. Note that

$$\rho(\mathbf{v}) = r_{\mathbf{w}}(r_{\mathbf{v}}(\mathbf{v})) = -(r_{\mathbf{w}}(\mathbf{v})) = -(\mathbf{v} - 2\pi(\mathbf{v})) = 2\pi(\mathbf{v}) - \mathbf{v}$$

where $\pi = \pi_{\mathbf{w}}$. Checking the formula for $\pi(\mathbf{v})$ we have $\mathbf{v}\bullet\pi(\mathbf{v}) = \frac{(\mathbf{v}\bullet\mathbf{w})^2}{\mathbf{w}\bullet\mathbf{w}}$ therefore

$$\rho(\mathbf{v})\bullet\mathbf{v} = 2\frac{(\mathbf{v}\bullet\mathbf{w})^2}{\mathbf{w}\bullet\mathbf{w}} - \mathbf{v}\bullet\mathbf{v}.$$

Since ρ is orthogonal, we also have that $\|\rho(\mathbf{v})\| = \|\mathbf{v}\|$. Therefore

$$\cos\theta(\rho(\mathbf{v}),\mathbf{v}) = \frac{\rho(\mathbf{v})\bullet\mathbf{v}}{\|\mathbf{v}\|^2} = 2\frac{(\mathbf{v}\bullet\mathbf{w})^2}{(\mathbf{w}\bullet\mathbf{w})(\mathbf{v}\bullet\mathbf{v})} - 1 = 2\left(\frac{(\mathbf{v}\bullet\mathbf{w})^2}{\|\mathbf{w}\|\cdot\|\mathbf{v}\bullet\mathbf{v}\|}\right)^2 - 1 = 2\cos^2\theta - 1 = \cos 2\theta.$$

Therefore $\alpha = 2\theta$.

(4) Lastly, we need to show that $\epsilon(\mathbf{v}\times\mathbf{w}) = +1$. Indeed,

$$\det(\mathbf{v},\rho(\mathbf{v}),\mathbf{u}) = \det(\mathbf{v},2\pi(\mathbf{v})-\mathbf{v},\mathbf{u}) = 2\det(\mathbf{v},\pi(\mathbf{v}),\mathbf{u}) \sim (\mathbf{v}\bullet\mathbf{w})\cdot\det(\mathbf{v},\mathbf{w},\mathbf{u})$$

which is positive under the assumption that the angle between $\mathbf{v}$ and $\mathbf{w}$ is acute.

The proof is complete. ■

C.6 Rigid Motions

Proposition C.17. *The matrix representation of a rotation is an element of* $\mathbf{SO}_n(\mathbb{R})$.

Proof. Every rotation is the composition of two reflections. In particular they (rotations) are the composition of linear maps (reflections). Hence, by Example/Theorem 3.179 they are also linear. Since the determinant of any reflection is -1, we have that the product of any reflections must have determinant $-1\cdot -1 = 1$ by Theorem 3.205. ■

In dimensions 2 and 3, the remarkable thing is that the converse is also true.

Theorem C.18 (Euler's Theorem for Rotations)**.** $\mathbf{SO}_3(\mathbb{R})$ *is the set of 3 dimensional rotation matrices.*

Remark C.19. In dimension higher than 3 this theorem is false. For instance in dimension 4, negative the identity matrix,

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

is an element of $\mathbf{SO}_4(\mathbb{R})$ but is not the matrix of any rotation for any reasonable defintion of what a rotation would be in 4 dimensions (a rotation should certainly fix at least something besides the origin!).

Before proving the theorem we first prove a couple of lemmas:

Lemma C.20. *Suppose that* $\mathbf{u}$ *and* $\mathbf{v}$ *are two vectors in* $\mathbb{R}^3$ *which are the same distance from the origin. Then the reflection* $r_{\mathbf{u}-\mathbf{v}}$ *switches* $\mathbf{u}$ *and* $\mathbf{v}$ (*in the sense that* $r_{\mathbf{u}-\mathbf{v}}(\mathbf{u}) = \mathbf{v}$ *and vice-versa.*)

Proof. Let c denote the number $\|\mathbf{u}\|^2$. Then we have that

$$\mathbf{u} \bullet (\mathbf{u} - \mathbf{v}) = c - \mathbf{u} \bullet \mathbf{w} \quad \text{and} \quad (\mathbf{u} - \mathbf{w}) \bullet (\mathbf{u} - \mathbf{w}) = 2c - 2\mathbf{u} \bullet \mathbf{v} = 2(c - \mathbf{u} \bullet \mathbf{v}).$$

Hence

$$\begin{aligned} r_{\mathbf{u}-\mathbf{v}}(\mathbf{u}) &= \mathbf{u} - 2\frac{\mathbf{u} \bullet (\mathbf{u} - \mathbf{v})}{(\mathbf{u} - \mathbf{v}) \bullet (\mathbf{u} - \mathbf{v})}(\mathbf{u} - \mathbf{v}) \\ &= \mathbf{u} - \frac{2(c - \mathbf{u} \bullet \mathbf{w})}{2(c - \mathbf{u} \bullet \mathbf{v})}(\mathbf{u} - \mathbf{v}) = \mathbf{u} - \mathbf{u} + \mathbf{v} = \mathbf{v}. \end{aligned}$$

And since any reflection is an involution, we have that $r_{\mathbf{u}-\mathbf{v}}(\mathbf{v}) = \mathbf{u}$ as well. ∎

Definition C.21. A function $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$ is a *rigid motion* or *isometry* if it preserves the distance between any two points. That is for all $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$

$$d(\mathbf{v}, \mathbf{w}) = d(\varphi(\mathbf{v}), \varphi(\mathbf{w})).$$

The set of all isometries of $\mathbb{R}^n$ is denoted $\operatorname{Isom}\mathbb{R}^n$ and is called the *isometry group* of $\mathbb{R}^n$.

The precise definition of what we've been calling the symmetries of X, which we've been denoting as $S^*(X)$ is really the set of isometries which take a set X to itself. What is the rigorous definition of $S(X)$? Well, it's the set of orientation preserving isometries which preserve X.

For instance, if $n = 2$, you should picture a rigid motion as something which is possible if the plane is made out of plywood.

Example C.22. Every orthogonal transformation is a rigid motion. In addition, every translation by any vector is a rigid motion since if $\mathbf{u}$ is a vector, then translation by

$\mathbf{u}$, $\tau_{\mathbf{u}}$, has the formula $\tau_{\mathbf{u}}(\mathbf{v}) = \mathbf{v} + \mathbf{u}$. Therefore,

$$\begin{aligned} d(\tau_{\mathbf{u}}(\mathbf{v}), \tau_{\mathbf{u}}(\mathbf{w})) &= \| \tau_{\mathbf{u}}(\mathbf{v}) - \tau_{\mathbf{u}}(\mathbf{w}) \| \\ &= \| (\mathbf{v} + \mathbf{u}) - (\mathbf{w} + \mathbf{u}) \| \\ &= \| \mathbf{v} - \mathbf{w} \| \\ &= d(\mathbf{v}, \mathbf{w}). \end{aligned}$$

(This should certainly be intuitively clear). However, since translations do not take the origin to the origin (unless $\mathbf{v} = \mathbf{0}$) they are not linear transformations. In a sense, we will see that these are the only family of rigid motions which are not linear transformations in the sense that if φ is a rigid motion, then there exists a unique orthogonal (hence linear) transformation, $\mathbf{A}$, and a unique vector $\mathbf{v}$ so that $\varphi = \tau_{\mathbf{v}} \circ \mathbf{A}$.

EXERCISE C.7. Suppose that φ and φ' are rigid motions. Prove that $\varphi \circ \varphi'$ is also a rigid motion.

SOLUTION

The following lemma would be immediate if we knew before hand that rigid motions fixing the origin were linear, but as we don't already know this we have to put in a bit of effort!

LEMMA C.23. *If a rigid motion of $\mathbb{R}^n$ which fixes all standard basis vectors is the identity.*

PROOF. Note that for all $\mathbf{w} \in \mathbb{R}^n$ that

$$d(\mathbf{w}, \mathbf{v}) = \| \mathbf{w} - \mathbf{v} \|^2 = \| \mathbf{w} \|^2 - 2\mathbf{w} \bullet \mathbf{v} + \| \mathbf{v} \|^2 = d(\mathbf{w}, \mathbf{0})^2 - 2\mathbf{w} \bullet \mathbf{v} + d(\mathbf{v}, \mathbf{0})^2.$$

Therefore, if we suppose φ fixes the standard basis vector $\mathbf{e}_1$. Then since for all $\mathbf{v} \in \mathbb{R}^n$

$$d(\varphi(\mathbf{v}), \mathbf{e}_i) = d(\varphi(\mathbf{v}), \varphi(\mathbf{e}_i)) = d(\mathbf{v}, \mathbf{e}_i)$$

we see that the two quantities

$$d(\varphi(\mathbf{v}), \mathbf{e}_i)^2 = d(\varphi(\mathbf{v}), \mathbf{0})^2 - 2\varphi(\mathbf{v}) \bullet \mathbf{e}_i + d(\mathbf{e}_i, \mathbf{0})^2 = d(\mathbf{v}, \mathbf{0})^2 - 2\varphi(\mathbf{v}) \bullet \mathbf{e}_i + 1$$

and

$$d(\mathbf{v}, \mathbf{e}_i)^2 = d(\mathbf{v}, \mathbf{0})^2 - 2\mathbf{v} \bullet \mathbf{e}_i + 1$$

are equal. Therefore $(\varphi(\mathbf{v}) - \mathbf{v}) \bullet \mathbf{e}_i = 0$ for all i. But this is the i^{th} coordinate of the difference $(\varphi(\mathbf{v}) - \mathbf{v})$. Hence all coordinates of this difference are zero, i.e. this difference is zero, i.e. $\varphi(\mathbf{v}) = \mathbf{v}$. ■

THEOREM C.24. *Every*[I] *element of* $\text{Isom}^\circ(\mathbb{R}^n)$ *can be written as the product of m reflections where m is some positive integer less than or equal to n.*

PROOF. Let $\mathbf{A} \in \text{Isom}^\circ(\mathbb{R}^n)$ be arbitrary. We assume that $\mathbf{A}$ is not the identity (see the footnote below). This proof actually gives an algorithm to follow. We state this algorithm it inductively.

(1) Define $\mathbf{A}_0$ as $\mathbf{A}$ itself.

(2) Now assume that $\mathbf{A}_{i-1}$ has been defined. Define

$$\mathbf{A}_i := r_{\mathbf{v}} \circ \mathbf{A}_{i-1}$$

where $\mathbf{v} = \mathbf{A}_{i-1}(\mathbf{e}_i) - \mathbf{e}_i$ using the convention that $r_{\mathbf{0}} = \mathbf{1}_{\mathbb{R}^n}$.

[I] Just to set things up so they work out well, let's define the product of zero reflections to be the identity.

Since it'll come up we define $\mathbf{v}_i$ as this (possibly zero) vector $\mathbf{v}$ and $r_i := r_{\mathbf{v}_i}$. Note r_i isn't a reflection if $\mathbf{v}_i$ happens to be $\mathbf{0}$, but no matter we still have that $r_i^2 = \mathbf{I}$, and $r_i(\mathbf{e}_i) = \mathbf{A}_{i-1}(\mathbf{e}_i)$, where we're using Lemma C.20. Composing both sides of this second equation with r_i, and using the definition that $\mathbf{A}_i = r_i \circ \mathbf{A}_{i-1}$, we see that

$$\mathbf{A}_i(\mathbf{e}_i) = \mathbf{e}_i.$$

So $\mathbf{A}_i$ fixes $\mathbf{e}_i$. Moreover, each of these $\mathbf{A}_i$'s are, as the product of orthogonal transformation, orthogonal transformations. I claim that each $\mathbf{A}_i$, in addition to $\mathbf{e}_i$, also fixes each $\mathbf{e}_j$ for $1 \leq j \leq i$. Since there is no such thing as $\mathbf{e}_0$, this is vacuously true in the case that $i = 0$. Proceeding by induction, we therefore, assume that it's true for $i-1$. That is we are assuming that $\mathbf{A}_{i-1}(\mathbf{e}_j) = \mathbf{e}_j$ for every $1 \leq j \leq i-1$.

Let $j < i$, then we have

$$\mathbf{e}_i \bullet \mathbf{e}_j = \mathbf{A}_{i-1}(\mathbf{e}_i) \bullet \mathbf{A}_{i-1}(\mathbf{e}_j) = \mathbf{A}_{i-1}(\mathbf{e}_i) \bullet \mathbf{e}_j \Rightarrow \mathbf{A}_{i-1}(\mathbf{e}_i) \bullet \mathbf{e}_j - \mathbf{e}_i \bullet \mathbf{e}_j = 0 \Rightarrow (\mathbf{A}_{i-1}(\mathbf{e}_i) - \mathbf{e}_i) \bullet \mathbf{e}_j = 0 \Rightarrow (\mathbf{A}_{i-1}(\overset{\mathbf{v}_i}{\overset{\shortparallel}{\mathbf{e}_i}}) - \mathbf{e}_i) \perp \mathbf{e}_j.$$

Hence r_i, as a reflection in a line perpendicular to $\mathbf{e}_j$, fixes, $\mathbf{e}_j$. Therefore $\mathbf{A}_i$ is the composition of functions, r_i and $\mathbf{A}_{i-1}$ fixing $\mathbf{e}_j$, and hence $\mathbf{A}_i$ itself must fix $\mathbf{e}_j$, as we claimed.

In summary, for every orthogonal transformation $\mathbf{A}$, we have found a sequence of transformations,

$$r_1, \ldots, r_n,$$

which are either the identity or a (honest to goodness) reflection, with

$$\mathbf{A}_n = r_n \circ \cdots r_1 \circ \mathbf{A}.$$

Now $\mathbf{A}_n$is an element of $\mathrm{Isom}^\circ(\mathbb{R}^n)$ and, as we just proved, fixes all n standard basis vectors, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of $\mathbb{R}^n$. Therefore, by Lemma C.23, $\mathbf{A}_n = \mathbf{1}_{\mathbb{R}^n}$. So using $r_i^2 = \mathbf{1}$ to move each to the other side (going from left to right)

$$\mathbf{I}_n = r_n \circ \cdots r_1 \circ \mathbf{A} \Rightarrow r_1 \circ \cdots \circ r_n = \mathbf{A}.$$

Therefore, we have written $\mathbf{A}$ as the composition of n or fewer (fewer in the case that some of the r_i's were taken to be the identity) reflections. ■.

This proof actually gives an algorithm to follow. We state it inductively

(1) Define $\mathbf{A}_0$ as the original matrix $\mathbf{A}$.

(2) Now assume that $\mathbf{A}_{i-1}$ has been defined. If $\mathbf{A}_{i-1}$ fixes $\mathbf{e}_i$, i.e. if $\mathbf{A}_{i-1}(\mathbf{e}_i) = \mathbf{e}_i$, then define $\mathbf{A}_i$ as $\mathbf{A}_{i-1}$. On the other hand, if $\mathbf{A}_{i-1}$ does not fix $\mathbf{e}_i$, then define $\mathbf{A}_i$ as the matrix

$$r_{\mathbf{A}_{i-1}-\mathbf{e}_{i-1}} \cdot \mathbf{A}_{i-1}.$$

COROLLARY C.25. *For every rigid motion φ of $\mathbb{R}^n$ there is a product of m reflections ($m \leq n$) $\mathbf{R}$ so that*

$$\varphi = \tau \circ \mathbf{R}$$

where τ is translation in $\varphi(\mathbf{0})$, i.e. it's the map $\mathbf{v} \mapsto \mathbf{v} + \varphi(\mathbf{0})$. ■

In fact if you wish to consider reflections which aren't necessarily through the origin, it's easy to write any translation as the composition of exactly two reflections proving that every rigid motion is the product of at most $n+2$ reflections. Note also that while the reflections making up $\mathbf{R}$ are not necessarily unique, their product $\mathbf{R}$ as well as the translation τ are unique.

LEMMA C.26. *Every rigid motion is a bijection.*

PROOF. Let φ be a rigid motion and denote $\varphi(\mathbf{0})$ as $\mathbf{v}$. Then $\tau_{-\mathbf{v}} \circ \varphi$ is a rigid motion fixing the origin. Therefore it is the composition of reflections, $\mathbf{R}$. Therefore $\varphi = \tau_{\mathbf{v}} \circ \mathbf{R}$ and φ is the composition of bijections, hence a bijection. ■

It is not too much of a stretch to require that a rigid motion fix the origin since if φ is a rigid motion and $\mathbf{v} = \varphi(\mathbf{0})$, then define a new rigid motion $\varphi' = \tau_{-\mathbf{v}} \circ \varphi$. We denote the set of rigid motions which fix the origin as

$$\text{Isom}^\circ \mathbb{R}^n.$$

THEOREM C.27. *The set of rigid motions which fix the origin is* $\mathbf{O}_n(\mathbb{R})$. *That is*

$$\mathbf{O}_n(\mathbb{R}) = \text{Isom}^\circ \mathbb{R}^n.$$

That is rigid motions are orthogonal and linear.

PROOF.

$\subseteq$ Elements of $\mathbf{O}_n(\mathbb{R})$ are rigid motions which fix the origin.

$\supseteq$ Elements of $\text{Isom}^\circ \mathbb{R}^n$ are products of reflections, each of which are elements of $\mathbf{O}_n(\mathbb{R})$, therefore so is their composition.

■

COROLLARY C.28. *For every rigid motion* φ *there exists a unique translation* τ *and orthogonal transformation* $\mathbf{A}$ *so that*

$$\varphi = \tau \circ \mathbf{A}.$$

PROOF. As pointed out already, τ must be translation in $\varphi(\mathbf{0})$ and $\mathbf{A}$ must therefore be $\tau^{-1} \circ \varphi$.

■

Initially, for a student first trying to learn group theory, it's a bit dangerous to attempt to make this anymore precise than this intuitive concept, but later we will study the set of symmetries of an object in more detail.

We give a chapter 4 friendly proof of the following:

THEOREM C.29. *For any set* $X \subseteq \mathbb{R}^n$, *the set of symmetries of* X *and the set of op symmetries of* X *form a group under function composition.*

REMARK C.30. In other words for any set $X \subseteq \mathbb{R}^3$, $(S^*(X), \circ)$ and $(S(X), \circ)$ are groups.[II]

PROOF.

Closure Suppose that φ and ψ are symmetries of X.

- We need to show that $\varphi(\psi(X)) = X$. But this is easy since
$$(\varphi \circ \psi)(X) = \varphi(\psi(X)) = \varphi(X) = X.$$
- The composition is a rigid motion due to Exercise C.7.

Assoc. The elements of $S^*(X)$ are functions and composition of functions is associative in general.

Identity The identity function is clearly an element of $S^*(X)$.

Inverses Let $\varphi \in S^*(X)$. Then, by Lemma C.26, φ is a bijection so φ^{-1} exists. Moreover we already know that φ^{-1} is a bijection. We still need to show that φ^{-1} preserves X and distances.

- Since
$$\varphi(X) = X \Rightarrow \cancel{\varphi^{-1}(\varphi(X))} = \varphi^{-1}(X) \Rightarrow X = \varphi^{-1}(X)$$
and φ^{-1} preserves X.
- Let $x, y \in \mathbb{R}^3$. Then we have
$$d(x,y) = d(\varphi(\varphi^{-1}(x)), \varphi(\varphi^{-1}(y))) = d(\varphi^{-1}(x), \varphi^{-1}(y))$$
and φ^{-1} also preserves distances. ■

REMARK C.31. Notice that if an orthogonal transformation $\mathbf{A}$ is the composition of m reflections

$$\mathbf{A} = r_1 \circ \cdots \circ r_m,$$

then, again by Theorem 3.205, the determinant of $\mathbf{A}$ is $(-1)^m$ which is 1 if m is even and -1 if m is odd. Therefore one can characterize $\mathbf{SO}_n(\mathbb{R})$ as the set of orthogonal transformation which can be written as a product of an even number of reflections.

PROOF. (Proof of theorem C.18) The only even number between 1 and 3 is 2. Therefore every element of $\mathbf{SO}_3(\mathbb{R})$ can be written as a product of two reflections, $r_{\mathbf{u}}$ and $r_{\mathbf{v}}$. If $\mathbf{u}$ is proportional to $\mathbf{v}$ their product is the identity, if they are not proportional their product is a rotation. ■

EXAMPLE C.32. The matrix

$$\mathbf{A} := \frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \end{pmatrix}$$

is a rotation matrix. Indeed we have that

$$\begin{aligned}
\mathbf{A}^\top \mathbf{A} &= \frac{1}{9}\begin{pmatrix} 1 & 2 & 2 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \end{pmatrix}\begin{pmatrix} 1 & -2 & -2 \\ 2 & 2 & -1 \\ 2 & -1 & 2 \end{pmatrix} \\
&= \frac{1}{9}\begin{pmatrix} 1+4+4 & 2-4+2 & 2+2-4 \\ -4+4+1 & 4+4+1 & 4-2-2 \\ 2+2-4 & 4-2-2 & 4+1+4 \end{pmatrix} \\
&= \frac{1}{9}\begin{pmatrix} 9 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 9 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \mathbf{I}_3
\end{aligned}$$

and just looking at the $\mathbf{k}$ component of the cross product of the first two columns:

$$\left(\frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix} \times \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}\right) \bullet k = \frac{1}{9}(1 \cdot 2 - 2 \cdot (-2)) = \frac{2}{3}$$

which is the **k** component of the third column, we can see automatically that **A** has determinant 1 (why?) and so, by definition, is an element of $\mathbf{SO}_3(\mathbb{R})$.

Therefore, by definition, $\mathbf{A} \in \mathbf{SO}_3(\mathbb{R})$, which Euler's theorem says is the matrix of some rotation.

In fact the proof of Euler's theorem gives us an algorithm to write this rotation as the product of two reflections. Since (only using angled brackets here (and only here) instead of parenthesis here because it looks a bit cleaner with so many parenthesis running around already!)

$$\mathbf{A}(\mathbf{e}_1) = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix}$$

we have

$$\mathbf{e}_1 - \mathbf{A}(\mathbf{e}_1) = \frac{2}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Since reflections ignore the scalar multiple, we scale this vector (for ease) and define

$$\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

By construction,

$$r_{\mathbf{v}}(\mathbf{e}_1) = \begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix}.$$

Moreover,

$$r_{\mathbf{v}}(\mathbf{e}_2) = \mathbf{e}_2 - 2 \cdot \frac{\mathbf{e}_2 \bullet \mathbf{v}}{\mathbf{v} \bullet \mathbf{v}}\mathbf{v} = \mathbf{e}_2 - 2\frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3} \cdot \begin{bmatrix} -2 \\ 1 \\ -2 \end{bmatrix}$$

$$r_{\mathbf{v}}(\mathbf{e}_3) = \mathbf{e}_3 - 2 \cdot \frac{\mathbf{e}_3 \bullet \mathbf{v}}{\mathbf{v} \bullet \mathbf{v}}\mathbf{v} = \frac{1}{3} \cdot \begin{bmatrix} -2 \\ -2 \\ 1 \end{bmatrix}.$$

Therefore the matrix for $r_{\mathbf{v}}$ is

$$\frac{1}{3}\begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}.$$

From this we calculate that $r_{\mathbf{v}} \cdot \mathbf{A}$ is the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}.$$

(Of course there's only a single 1 in the first entry of the first column. This is what we wanted.) This is the matrix for $r_{\mathbf{e}_2+\mathbf{e}_3}$ Therefore we have "factored" the rotation **A** as the product of two reflections:

$$\mathbf{A} = r_{\mathbf{e}_1+\mathbf{e}_2+\mathbf{e}_3} \circ r_{\mathbf{e}_2+\mathbf{e}_3}.$$

Note also that since

$$(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) \times (\mathbf{e}_2 + \mathbf{e}_3) = \mathbf{e}_1 \times (\mathbf{e}_2 + \mathbf{e}_3) = \mathbf{e}_3 - \mathbf{e}_2$$

the axis of $\mathbf{A}$ is determined by the line

$$\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

However, this could have been seen initially since

$$\mathbf{A} \cdot \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \end{pmatrix} \cdot \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \frac{1}{3}\left(-\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix} \right) = \frac{1}{3}\begin{bmatrix} 0 \\ -3 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

EXERCISE C.8. Show that two reflections $r_{\mathbf{v}}$ and $r_{\mathbf{w}}$ commute iff $\mathbf{v}$ and $\mathbf{w}$ are either perpendicular or parallel.

SOLUTION

C.7 CONNECTION WITH QUATERNIONS

DEFINITION C.33. We define a map, $\mathbf{R}$ (we will end up showing $\mathbf{R}$ has image $\mathbf{SO}_3(\mathbb{R})$) by sending $\mathfrak{q} \in \mathbb{H}^\times$ to

$$\begin{aligned} \mathbf{R}(\mathfrak{q}) &: \mathbb{R}^3 \to \mathbb{R}^3 \\ &: \mathbf{v} \mapsto \mathfrak{q}\mathbf{v}\mathfrak{q}^{-1}. \end{aligned}$$

Typically we denote $\mathbf{R}(\mathfrak{q})$ as $\mathbf{R}_{\mathfrak{q}}$.

In the language of Chapter 10 on group actions, the map $\mathbf{R}$ is an action of, what we'll end up showing, $\mathbb{S}^3$ on $\mathbb{R}^3$ by isometries.

RANT C.34.

(1) For an arbitrary $\mathfrak{q} \in \mathbb{H}^\times$ (i.e. a non-zero quaternion) we write $\mathfrak{q}$ as $\mathfrak{q} = c\widehat{\mathfrak{q}}$ where $\widehat{\mathfrak{q}} \in \mathbb{S}^3$ (c happens to equal $\|\mathfrak{q}\|$ and $\widehat{\mathfrak{q}}$ happens to equal $\mathfrak{q}/\|\mathfrak{q}\|$). Then

$$\begin{aligned} \mathbf{R}_{\mathfrak{q}}(\mathbf{v}) &= (c\widehat{\mathfrak{q}})\mathbf{v}(c\widehat{\mathfrak{q}})^{-1} \\ &= c\widehat{\mathfrak{q}}\mathbf{v}c^{-1}\widehat{\mathfrak{q}}^{-1} \\ &= cc^{-1}\widehat{\mathfrak{q}}\mathbf{v}\widehat{\mathfrak{q}}^{-1} \\ &= \widehat{\mathfrak{q}}\mathbf{v}\widehat{\mathfrak{q}}^{-1} \\ &= \mathbf{R}_{\widehat{\mathfrak{q}}}(\mathbf{v}), \end{aligned}$$

i.e. $\mathbf{R}_{\mathfrak{q}} = \mathbf{R}_{\widehat{\mathfrak{q}}}$. As such, we will only really need to consider quaternions in $\mathbb{S}^3$, i.e. we take the domain of $\mathbf{R}$ to be $\mathbb{S}^3$ instead of all of $\mathbb{H}^\times$.

(2) All of the following will be implied by Theorem C.35 and we will not need to prove them directly: $\mathbf{R}_q$ is

- A map $\mathbb{R}^3 \to \mathbb{R}^3$.
- Linear.
- Orthogonal, i.e. an element of $\mathbf{SO}_3(\mathbb{R})$.

Even though we do not need to prove them directly, they all can be readily proved in this way. Perhaps it will be worth it for you to do just this.

EXERCISE C.9. Prove directly from the definition that $\mathbf{R}_q$ is linear.

SOLUTION

EXERCISE C.10. Prove directly from the definition that $\mathbf{R}_q$ maps $\mathbb{R}^3$ to $\mathbb{R}^3$ (i.e. takes pure quaternions to pure quaternions).

SOLUTION

EXERCISE C.11. Prove directly from the definition that $\mathbf{R}_q$ is orthogonal.
Hint! Just need to prove that for any $\mathbf{v} \in \mathbb{R}^3$, $\left\| \mathbf{R}_q(\mathbf{v}) \right\| = \|\mathbf{v}\|$.

SOLUTION

The main goal is to prove:

THEOREM C.35. *Suppose that*

$$q = e^{\frac{\theta}{2}\widehat{\mathbf{u}}} \Rightarrow \mathbf{R}_q = \varrho_{\theta,\widehat{\mathbf{u}}}.$$

In other words we have for all $\mathbf{v} \in \mathbb{R}^3$

$$q\mathbf{v}q^{-1} = \varrho_{\theta,\widehat{\mathbf{u}}}(\mathbf{v}).$$

Note that this even makes sense for the two elements of $\mathbb{S}^3$ without an axis, namely 1 and -1 since both of these go to the identity map.

For variety, we present two rather different proof of this theorem. (Of course, we need to only present one, but as usual the *pudding is in the proof*!)

C.7.1 PROOF 1

LEMMA C.36. *Suppose that* $\widehat{\mathbf{u}} \in \mathbb{S}^2$. *Then for any* $\mathbf{v} \in \mathbb{R}^3$ *we have that*

$$\widehat{\mathbf{u}} * \mathbf{v} * \widehat{\mathbf{u}} = r_{\widehat{\mathbf{u}}}(\mathbf{v})$$

where the product on the LHS is quaternion multiplication.

REMARK C.37. For $\widehat{\mathbf{u}} \in \mathbb{S}^2$, we have $-\widehat{\mathbf{u}} = \widehat{\mathbf{u}}^{-1}$, we are really proving that

$$-\mathbf{R}_{\widehat{\mathbf{u}}} = r_{\widehat{\mathbf{u}}}.$$

Proof. As is the general case multiplication of pure quaternions, we have for $\mathbf{v}$ and $\widehat{\mathbf{u}}$ that

$$\mathbf{v} * \widehat{\mathbf{u}} + \widehat{\mathbf{u}} * \mathbf{v} = -2\mathbf{v} \bullet \widehat{\mathbf{u}} \Rightarrow \mathbf{v} * \widehat{\mathbf{u}} = -2\mathbf{v} \bullet \widehat{\mathbf{u}} - \widehat{\mathbf{u}} * \mathbf{v}.$$

Hence

$$\begin{aligned}
\widehat{\mathbf{u}} * \mathbf{v} * \widehat{\mathbf{u}} &= \widehat{\mathbf{u}} * (-2\mathbf{v} \bullet \widehat{\mathbf{u}} - \widehat{\mathbf{u}} * \mathbf{v}) \\
&= -(\widehat{\mathbf{u}} * \widehat{\mathbf{u}}) * \mathbf{v} - 2(\mathbf{v} \bullet \widehat{\mathbf{u}})\widehat{\mathbf{u}} \\
&= \mathbf{v} - 2(\mathbf{v} \bullet \widehat{\mathbf{u}})\widehat{\mathbf{u}} && (\widehat{\mathbf{u}} \in \mathbb{S}^2 \Rightarrow \widehat{\mathbf{u}}^2 = -1) \\
&= r_{\widehat{\mathbf{u}}}(\mathbf{v}). && \text{(Equation (2))}
\end{aligned}$$

■

Proof. (**Proof of Theorem C.35, first go**) By Proposition 3.114 there are $\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2 \in \mathbb{S}^2$ with

$$\mathsf{q} = \widehat{\mathbf{u}}_1 * \widehat{\mathbf{u}}_2.$$

In this case we must have that

$$\widehat{\mathbf{u}}_1 \times \widehat{\mathbf{u}}_2 = \sin\theta\widehat{\mathbf{u}}.$$

Therefore

$$r_{\widehat{\mathbf{u}}_1} \circ r_{\widehat{\mathbf{u}}_2}$$

must be a rotation about $\widehat{\mathbf{u}}$ by an angle 2θ, i.e. it is equal to $\varrho_{2\theta,\widehat{\mathbf{u}}}(\mathbf{v})$. But by the lemma, we know that

$$\begin{aligned}
r_{\widehat{\mathbf{u}}_1} \circ r_{\widehat{\mathbf{u}}_2}(\mathbf{v}) &= \widehat{\mathbf{u}}_1 * (\widehat{\mathbf{u}}_2 * \mathbf{v} * \widehat{\mathbf{u}}_2) * \widehat{\mathbf{u}}_1 \\
&= \widehat{\mathbf{u}}_1 * (\widehat{\mathbf{u}}_2 * \mathbf{v} * -\widehat{\mathbf{u}}_2) * -\widehat{\mathbf{u}}_1 \\
&= \widehat{\mathbf{u}}_1 * (\widehat{\mathbf{u}}_2 * \mathbf{v} * \overline{\widehat{\mathbf{u}}_2}) * \overline{\widehat{\mathbf{u}}_1} \\
&= (\widehat{\mathbf{u}}_1 * \widehat{\mathbf{u}}_2) * \mathbf{v} * \overline{(\widehat{\mathbf{u}}_1 * \widehat{\mathbf{u}}_2)} \\
&= \mathsf{q} * \mathbf{v} * \overline{\mathsf{q}} \\
&= \mathsf{q} * \mathbf{v} * \mathsf{q}^{-1}.
\end{aligned}$$

■

C.7.2 Proof 2

Lemma C.38. *Let* $\mathsf{q} \in \mathbb{H}^\times$. *Suppose that* $\mathbf{v} \asymp \mathbf{Pu}\,\mathsf{q}$. *Then*

$$\mathbf{R}_{\mathsf{q}}(\mathbf{v}) = \mathbf{v}.$$

Proof. Since $\mathbf{v} \asymp \mathbf{Pu}\,\mathsf{q}$ we have that q and $\mathbf{v}$ commute. Hence

$$\mathbf{R}_{\mathsf{q}}(\mathbf{v}) = \mathsf{q}\mathbf{v}\mathsf{q}^{-1} = \mathbf{v}$$

as promised. ■

Lemma C.39. *Let* $\mathsf{q} = e^{\frac{\theta}{2},\widehat{\mathbf{u}}} \in \mathbb{S}^3$ *and suppose* $\mathbf{v} \perp \mathbf{Pu}\,\mathsf{q}$[1], *then*

$$\mathbf{R}_{\mathsf{q}}(\mathbf{v}) = \varrho_{\theta,\widehat{\mathbf{u}}}(\mathbf{v}).$$

[1] That is $\mathbf{v} \asymp \widehat{\mathbf{u}}$

Proof. We have

$$\begin{aligned}
\mathbf{R}_q(\mathbf{v}) &= q\mathbf{v}q^{-1} && \text{(Def'n of } \mathbf{R}\text{)}\\
&= qq\mathbf{v} && \text{(Since } \mathbf{v} \perp \widehat{\mathbf{u}}\text{)}\\
&= q^2\mathbf{v}\\
&= e^{\theta,\widehat{\mathbf{u}}}\mathbf{v}\\
&= (\cos\theta + \sin\theta\,\widehat{\mathbf{u}})\ \mathbf{v} && \text{(Def'n of } e^{\theta,\widehat{\mathbf{u}}}\text{)}\\
&= (\cos\theta\,\mathbf{v} + \sin\theta\,\widehat{\mathbf{u}} * \mathbf{v}) && \text{(Distrib)}\\
&= \cos\theta\,\mathbf{v} + \sin\theta\,\widehat{\mathbf{u}} \times \mathbf{v} && (\mathbf{v} * \widehat{\mathbf{u}} = \cancelto{0}{-\mathbf{v} \bullet \widehat{\mathbf{u}}} + \widehat{\mathbf{u}} \times \mathbf{v})\\
&= \varrho_{\theta,\widehat{\mathbf{u}}}(\mathbf{v}). && \text{(Equation (5))}
\end{aligned}$$

■

Proof. (**Proof of Theorem C.35, second go**) For an arbitrary $\mathbf{v} \in \mathbb{R}^3$ break $\mathbf{v}$ into its projections onto $\widehat{\mathbf{u}}$ and $\widehat{\mathbf{u}}^\perp$, i.e.

$$\mathbf{v} = \pi_{\widehat{\mathbf{u}}}(\mathbf{v}) + \pi_{\widehat{\mathbf{u}}^\perp}(\mathbf{v}),$$

then apply $\mathbf{R}_q$ while using both lemmas. ■

This exercise should feel a bit off at first. If you were to guess, you'd guess that the negative sign before the second θ does not belong there. Alas, it does.

Exercise C.12. Prove that for any reflection $\mathbf{r}$

$$\mathbf{r} \circ \varrho_{\theta,\widehat{\mathbf{u}}} \circ \mathbf{r}^{-1} = \varrho_{-\theta,\mathbf{r}(\widehat{\mathbf{u}})}.$$

Solution

For example, if $\mathbf{r} = r_{\mathbf{v}}$ where $\mathbf{v} \perp \widehat{\mathbf{u}}$, then $\mathbf{r}(\widehat{\mathbf{u}}) = \widehat{\mathbf{u}}$ proving that

$$\mathbf{r} \circ \varrho_{\theta,\widehat{\mathbf{u}}} \circ \mathbf{r}^{-1} = \varrho_{-\theta,\widehat{\mathbf{u}}} = \varrho^{-1}_{\theta,\widehat{\mathbf{u}}}.$$

This is always the case reflections and rotations (whose axis is $\mathbf{k}$) in the plane. This proves

Corollary C.40. *For any reflection* $\mathbf{r}$ *and rotation* ρ *in the plane we have*

$$\mathbf{r}\rho\mathbf{r}^{-1} = \rho^{-1}.$$

■

The intuition behind this corollary is that when you perform a rotation while looking in a mirror (the conjugation by the reflection "puts you in the mirror") you end up rotating in the opposite direction (hence the inverse).

C.7.3 General Results

In the language of Chapter 8 we can summarize our results:

Theorem C.41. *The map*

$$\begin{aligned}
\mathbf{R}\colon \mathbb{S}^3 &\to \mathbf{SO}_3(\mathbb{R})\\
: \cos\frac{\theta}{2} + \sin\frac{\theta}{2}\,\widehat{\mathbf{u}} &\mapsto \varrho_{\theta,\widehat{\mathbf{u}}}
\end{aligned}$$

is a surjective homomorphism (an epimorphism) with kernel $\{+1, -1\}$.

Proof. The map is surjective, since every element of $\mathbf{SO}_3(\mathbb{R})$ as a rotation, has the form $\varrho_{\theta,\widehat{\mathbf{u}}}$. Moreover the map is a homomorphism since $\mathbf{R}_{\mathsf{q}*\mathsf{p}} = \mathbf{R}_{\mathsf{q}} \circ \mathbf{R}_{\mathsf{p}}$ holds since $\mathbf{R}_{\mathsf{q}}$ is conjugation.

We can either see the fact that the kernel contains only 2 elements algebraically or geometrically. (I show both because I rock. Feel free to only pay attention to the proof you like better.)

- Geometrically, if $\varrho_{\theta,\widehat{\mathbf{u}}} = \mathbf{1}$ then we must have that θ is an integer multiple of 2π, $\theta = 2\pi n$ for some integer n. Hence the quaternion we started with

$$\cos\frac{\theta}{2} + \sin\frac{\theta}{2}\,\widehat{\mathbf{u}} = \cos\pi n + \sin\pi n\,\widehat{\mathbf{u}} = (-1)^n + 0 = (-1)^n.$$

- Algebraically, the kernel consists of the the set of all quaternions $\mathsf{q} \in \mathbb{S}^3$ which satisfy

$$\mathsf{q}\mathbf{v}\mathsf{q}^{-1} = \mathbf{v}$$

for all $\mathbf{v} \in \mathbb{R}^3$. In other words the kernel consists of the quaternions which commute with all pure quaternions. However this latter is obviously just $\{+1, -1\}$.[II] ■

Example C.42. Consider the cube, ◆[3] in $\mathbb{R}^3$ situated in the usual way (see 1 coming up on page 800 in the next appendix). We can rotate a cube by the either 90° or by −90° about each of its 6 faces. Since rotating by a given angle along a given face is the same as rotating by the opposite angle about the opposite face this gives rise to a total of

$$\frac{6\cdot 2}{2} = 6$$

which give rise to twice this number, or 12 quaternions, which we break up into 3 groups:

$$\pm\left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{i}\right) \qquad \pm\left(\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\mathbf{i}\right)$$
$$\pm\left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{j}\right) \qquad \pm\left(\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\mathbf{j}\right)$$
$$\pm\left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{k}\right) \qquad \pm\left(\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\mathbf{k}\right)$$

We can see easily that whenever we multiply together two quaternions of this form from different groupings we always get a quaternion of the form

$$\pm\frac{1}{2} \pm \frac{1}{2}\,\mathbf{i} \pm \frac{1}{2}\,\mathbf{j} \pm \frac{1}{2}\,\mathbf{k}.$$

For instance we have

$$\left(\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\mathbf{k}\right)\left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\mathbf{j}\right) = \frac{1}{2} + \frac{1}{2}\,\mathbf{i} + \frac{1}{2}\,\mathbf{j} - \frac{1}{2}\,\mathbf{k}. \tag{6}$$

All of this corresponds to the (in my opinion) less obvious fact that the composition of 90° rotations about distinct faces always leads to a 120° rotation about some vertex. For instance, the previous calculation is seen visually in Figure C.4.

[II] Actually recall that the only quaternions which commute with any pair non-proportional quaternions are real. Clearly there exist two non-proportional pure quaternions, for instance **i** and **j** do the job.

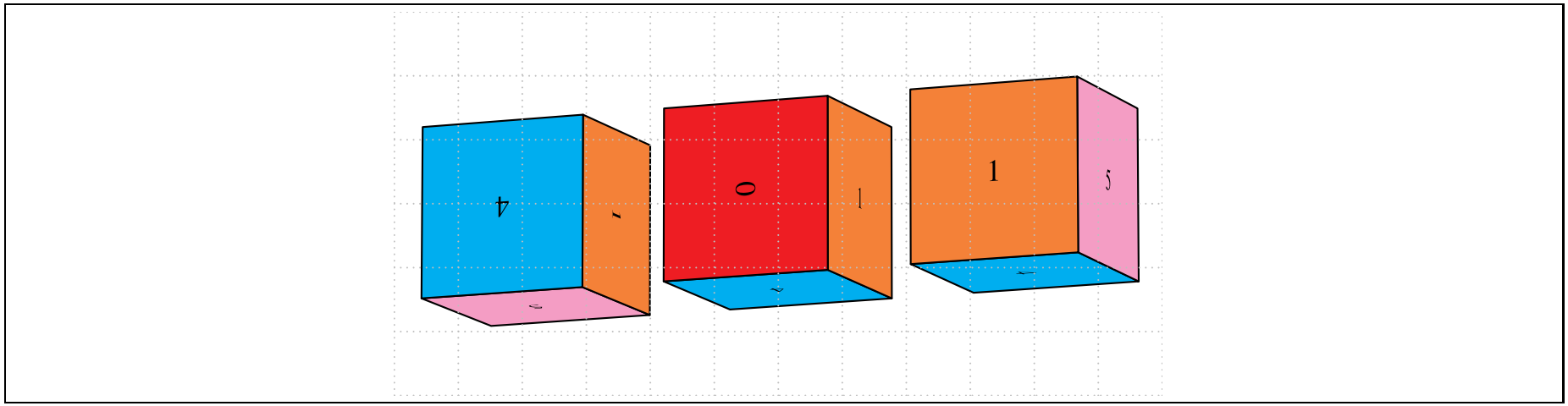


Figure C.4: Rotating a Cube

To go from the first picture to the second we performed the rotation $\mathcal{Q}_{\frac{\pi}{2},\,\mathbf{j}}$ and to get from the second to the third we perform the rotation $\mathcal{Q}_{\pi/2,\,-\mathbf{k}}$. From the figure we see that the net effect (to go from the first to the third picture) is $\mathcal{Q}_{\frac{\pi}{3},\,\mathbf{i}+\mathbf{j}-\mathbf{k}}$. Indeed, note the vertex in the bottom right hand corner, which is at the point $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, is fixed and the three faces which are colored pink, orange and blue are permuted in a counter-clockwise manner (from the viewpoint of $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$). Symbolically

$$\mathcal{Q}_{\frac{\pi}{2},\,-\mathbf{k}} \circ \mathcal{Q}_{\frac{\pi}{2},\,\mathbf{j}} = \mathcal{Q}_{\frac{\pi}{3},\,\mathbf{i}+\mathbf{j}-\mathbf{k}}. \tag{7}$$

which of course matches perfectly with our algebraic calculation in (6).

C.7.4 Explicit Matrix Representation.

To figure out an explicit matrix representation (with respect to the basis **i**,**j**, and **k**) takes a bit of annoying work. We work out the first part and leave the rest to you. Let $\mathsf{q} = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{S}^3$. Then $\mathsf{q}^{-1} = \overline{\mathsf{q}} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$.

$$\begin{aligned}
\mathsf{q}i\mathsf{q}^{-1} &= (a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k})i(a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}) \\
&= (a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k})(a - b\mathbf{i} + c\mathbf{j} + d\mathbf{k})\mathbf{i} \\
&= \Big(a^2 - (b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) \bullet (-b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) + (a \cdot (-b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) + a \cdot (b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) + (b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) \times (-b\mathbf{i} + c\mathbf{j} + d\mathbf{k})\Big)\mathbf{i} \\
&= \Big((a^2 + b^2 - c^2 - d^2) + (-ab\mathbf{i} + ac\mathbf{j} + ad\mathbf{k}) + (ab\mathbf{i} + ac\mathbf{j} + ad\mathbf{k}) + \cancel{(b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) \times (b\mathbf{i} + c\mathbf{j} + d\mathbf{k})} - (b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) \times 2b\mathbf{i}\Big)\mathbf{i} \\
&= \Big((a^2 + b^2 - c^2 - d^2) + (-ab\mathbf{i} + ac\mathbf{j} + ad\mathbf{k}) + (ab\mathbf{i} + ac\mathbf{j} + ad\mathbf{k}) + \mathbf{0} - (c\mathbf{j} + d\mathbf{k}) \times 2b\mathbf{i}\Big)\mathbf{i} \\
&= \Big((a^2 + b^2 - c^2 - d^2) + (-ab\mathbf{i} + ac\mathbf{j} + ad\mathbf{k}) + (ab\mathbf{i} + ac\mathbf{j} + ad\mathbf{k}) + (2bc\mathbf{k}) - (2bd\mathbf{j}))\Big)\mathbf{i} \\
&= \Big((a^2 + b^2 - c^2 - d^2) + (\cancel{-ab\mathbf{i}} + ac\mathbf{j} + ad\mathbf{k}) + (\cancel{ab\mathbf{i}} + ac\mathbf{j} + ad\mathbf{k}) + (2bc\mathbf{k}) - (2bd\mathbf{j}))\Big)\mathbf{i} \\
&= \Big((a^2 + b^2 - c^2 - d^2) + (2ad + 2bc)\mathbf{k} + (2ac - 2bd)\mathbf{j}\Big)\mathbf{i} \\
&= (a^2 + b^2 - c^2 - d^2)\mathbf{i} + (2ad + 2bc)\mathbf{j} - (2ac - 2bd)\mathbf{k} \\
&= (a^2 + b^2 - c^2 - d^2)\mathbf{i} + (2ad + 2bc)\mathbf{j} + (2bd - 2ac)\mathbf{k}.
\end{aligned}$$

A similiar calculation with **j** and then **k** replacing **i** shows us that

$$\begin{aligned}
\mathsf{q}i\mathsf{q}^{-1} &= (a^2 + b^2 - c^2 - d^2)\mathbf{i} + (2ad + 2bc)\mathbf{j} + (2bd - 2ac)\mathbf{k}. \\
\mathsf{q}j\mathsf{q}^{-1} &= (2bc - 2ad)\mathbf{i} + (a^2 - b^2 + c^2 - d^2)\mathbf{j} + (2cd + 2ab)\mathbf{k} \\
\mathsf{q}k\mathsf{q}^{-1} &= (2bd + 2ac)\mathbf{i} + (2cd - 2ab)\mathbf{j} + (a^2 - b^2 - c^2 + d^2)\mathbf{k}.
\end{aligned}$$

Now basic linear algebra says that the matrix of a linear transformation has columns equal to where the given basis goes. Therefore we have proven (for $\mathsf{q} = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{S}^3$)

Theorem C.43.

$$\mathbf{R}_{\mathsf{q}} = \begin{pmatrix} a^2 + b^2 - c^2 - d^2 & 2bc - 2ad & 2ac + 2bd \\ 2ad + 2bc & a^2 - b^2 + c^2 - d^2 & 2cd - 2ab \\ 2bd - 2ac & 2ab + 2cd & a^2 - b^2 - c^2 + d^2 \end{pmatrix}$$

■

Exercise C.13. Note the symmetry of the matrix: going from the i^{th} column to the $i + 1^{\text{th}}$ column (mod 3) permutes the rows according to the permutation $(1\ 2\ 3)$ while permuting the "variables" according to $(1)(b\ c\ d)^{-1} = (a)(b\ d\ c)$. Explain this phenomena as best as you can.

Hint! Consider the quaternion

$$\omega = \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}.$$

Recall that

$$\omega\mathbf{i}\omega^{-1} = \mathbf{j} \qquad \omega\mathbf{j}\omega^{-1} = \mathbf{k} \qquad \omega\mathbf{k}\omega^{-1} = \mathbf{i}.$$

Then, for $\mathsf{q} \in \mathbb{S}^3$ consider

$$\mathsf{q}j\mathsf{q}^{-1} = \mathsf{q}\omega\mathbf{i}\omega^{-1}\mathsf{q}^{-1}.$$

SOLUTION

Notice that if moreover q is pure, i.e. $\mathsf{q} \in \mathbb{S}^2$ this matrix simplifies a bit

COROLLARY C.44. *A* 180° *along an axis* $\widehat{\mathbf{u}} \in \mathbb{S}^2$ (*with* $\widehat{\mathbf{u}} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$) *is given by the matrix*

$$\varrho_{\pi,\widehat{\mathbf{u}}} = \begin{pmatrix} x^2 - y^2 - z^2 & 2xy & 2xz \\ 2xy & -x^2 + y^2 - z^2 & 2yz \\ 2xz & 2yz & -x^2 - y^2 + z^2 \end{pmatrix}$$

■

Note that this coincides with negative the reflection, $-r_{\widehat{\mathbf{u}}}$. If we were to plug in the spherical coordinates

$$x = \cos\varphi \quad y = \cos\theta \sin\varphi \quad z = \sin\theta \sin\varphi,$$

where θ is the latitude angle and φ is the longitudinal angle (treating $\begin{bmatrix}1\\0\\0\end{bmatrix}$ as the north pole), then, plugging these in for $\widehat{\mathbf{u}}$ shows $\varrho_{\pi,\widehat{\mathbf{u}}}$ takes the point $\begin{bmatrix}1\\0\\0\end{bmatrix} = \mathbf{i}$ to a point on $\mathbb{S}^2$ at the same latitude as $\widehat{\mathbf{u}}$, but at double the longitudinal angle of $\widehat{\mathbf{u}}$.

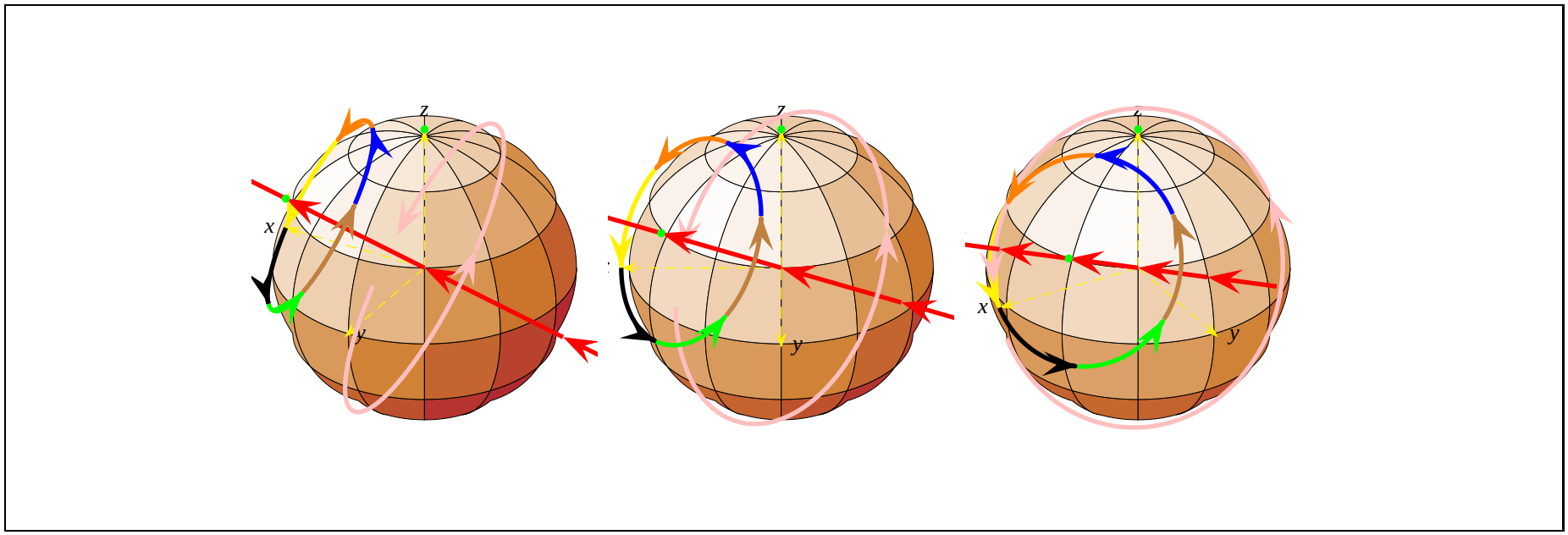


Figure C.5: Note that longitude of image of $\mathbf{i}$ is increased by angle twice the longitude of the axis.

The fact that this matrix is symmetric should be expected since these dudes represent 180° rotations, so they are orthogonal and of order 2, and for such a matrix

$$\mathbf{A} = \mathbf{A}^{-1} = \mathbf{A}^{\top},$$

i.e. $\mathbf{A}$ is symmetric. (See Exercise C.2 for more details.) Notice that this indeed is the matrix we'd get from first reflecting with line of reflection $\widehat{\mathbf{u}}$, and then multiplying by $-\mathbf{I}_3$.

EXAMPLE C.45. The matrix corresponding to

$$\widehat{\mathbf{u}} = \frac{\sqrt{2}}{2}(\mathbf{i} + \mathbf{j})$$

is

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

We consider the group of symmetries of the icosahedron instead of the dodecahedron. Place the icosahedron in front of us so that an edge lies perpendicular to the x–axis and so that the two endpoints of this edge lie in the $x - y$ plane. We find all the vertices of the icosahedron in this position. Let's define our starting vertex $\mathbf{v}_1$ to be the vertex which is the right endpoint of this starting edge and $\mathbf{v}_2$ be the left endpoint of this starting edge. Notice that a 180° rotation about the x–axis swaps $\mathbf{v}_1$ and $\mathbf{v}_2$. Since such a rotation is given by $\mathbf{R}_\mathbf{i}$, we have that

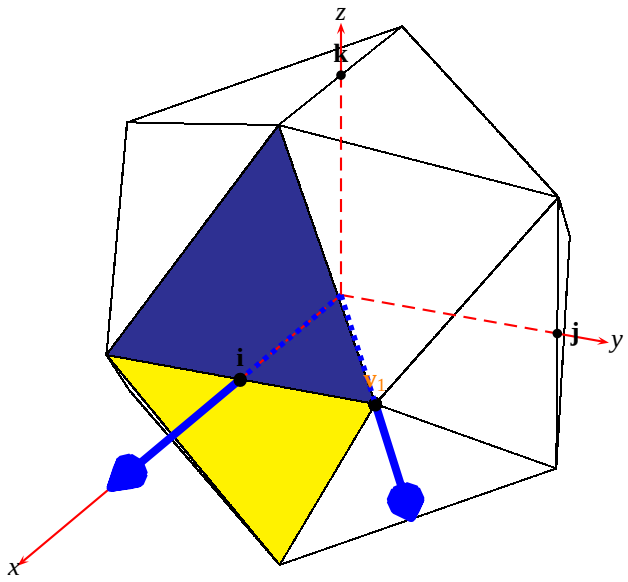
$$\mathbf{R}_\mathbf{i}(\mathbf{v}_1) = \varrho_{\pi,\mathbf{i}}(\mathbf{v}_1) = \mathbf{i}\,\mathbf{v}_1\,\mathbf{i}^{-1} = \mathbf{v}_2.$$

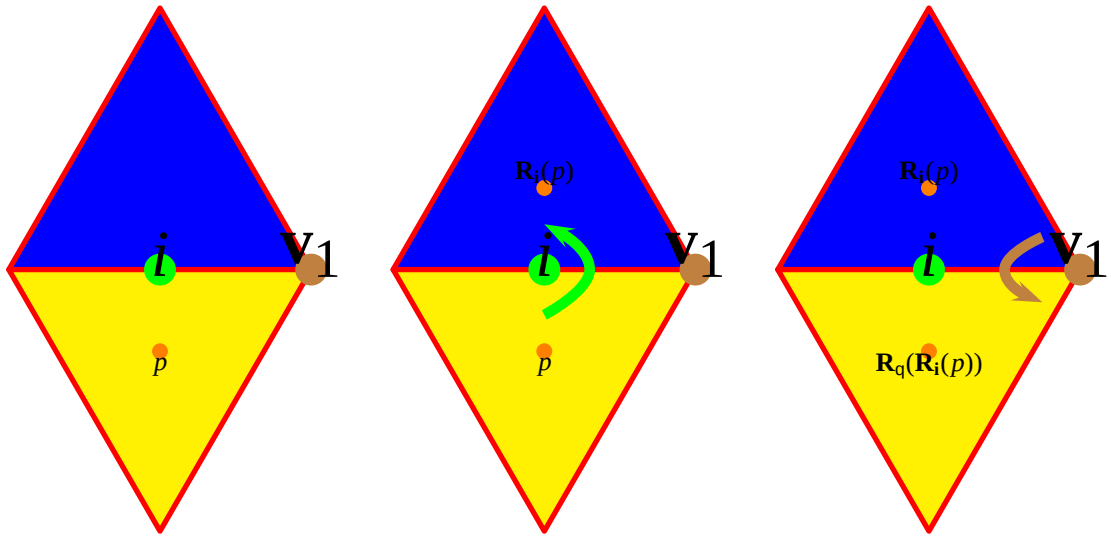
Clearly we have rotations of order five about every vertex, including rotation about our starting vertex $\mathbf{v}_1$. This is because each rotation about a vertex permutes the 5 adjacent vertices cyclically. Since $\mathbf{v}_1$ is in the x, y plane, the $\mathbf{k}$ component of $\mathbf{v}$ is zero. Thus rotation along $\mathbf{v}_1$ has the form

$$\mathfrak{q}_5 = \cos\frac{\pi}{5} + x\mathbf{i} + y\mathbf{j} = \frac{\phi}{2} + x\,\mathbf{i} + y\,\mathbf{j}$$

where ϕ is the golden ratio and $x\,\mathbf{i} + y\,\mathbf{j}$ is proportional to $\mathbf{v}_1$. We need to solve for the real numbers x and y.

Now if we actually play with an icosahedron, you can see that rotation along the front edge, which corresponds to the quaternion $\mathbf{i}$ (i.e. it equals to the rotation $\mathbf{R}_\mathbf{i}$) followed by the rotation corresponding to $\mathfrak{q}_5$ (i.e. it is equal to the rotation $\mathbf{R}_{\mathfrak{q}_5}$) has order 3. Indeed, the fastest way of seeing this is by noting that this composition fixes a face. We can actually seeing this simply by drawing two adjacent triangles of the icosahedron:





Rot 180° along edge, **i**

Rot 72° along vertex, $\mathbf{v}_1$

From the picture, we see the net effect fixes the vertex right below the edge centered at $\mathbf{i}$. Since faces are triangles this means that the rotation $\mathbf{R}_q \circ \mathbf{R}_\mathbf{i}$ has order 3. [III] As such, the rotation corresponding to $q_5 * \mathbf{i}$ is a rotation by angle $\frac{2\pi}{3}$ about some axis. This means that $\mathbf{Re}(q_5 * \mathbf{i}) = \pm\frac{1}{2}$ (since $\pm\frac{1}{2} = \pm\cos\frac{\pi}{3}$). We can assume, WLOG, that it is equal to the negative version, namely $= -\frac{1}{2}$ (we can always replace replace a quaternion by its negative since each give rise to the same rotation). On the other hand, if we do the multiplication $q_5 * \mathbf{i}$ out directly we see that

$$q_5 * \mathbf{i} = \left(\frac{\phi}{2} + x\,\mathbf{i} + y\,\mathbf{j}\right) * \mathbf{i} = -x + \frac{\phi}{2}\mathbf{i} - y\,\mathbf{k}.$$

Comparing with what we just saw from the picture, namely that $\mathbf{Re}(q_5 * \mathbf{i}) = -\frac{1}{2}$ we have solved for x, $x = \frac{1}{2}$[IV].

We still need to solve for the second, and final, unknown y. We still haven't used the fact that q_5 has unit length, i.e. that $q_5 \in \mathbb{S}^3$, and we do that now to also solve for our y.

$$\|q_5\|^2 = \left(\frac{\phi}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + y^2 = 1 \tag{8}$$

$$\Rightarrow \frac{1}{4} + \frac{\phi^2}{4} + y^2 = 1 \tag{9}$$

$$\Rightarrow 2 + \phi + 4y^2 = 4 \tag{10}$$

$$\Rightarrow 4y^2 = 2 - \phi. \tag{11}$$

Lucking the quantity, $2 - \phi$, on the right hand side of the resulting equation is a perfect square since a direct expansion gives us

$$(\phi - 1)^2 = \phi^2 - 2\phi + 1 = 1 + \phi - 2\phi + 1 = 2 - \phi.$$

We plug this back into (8)

$$\Rightarrow 4y^2 = 2 - \phi \Rightarrow y = \pm\frac{1}{2}(\phi - 1).$$

Indeed, since we took the vertex $\mathbf{v}_1$ the way we did (to the right of $\mathbf{i}$) we know that $y > 0$, and since $\phi > 1$, we take the root with the + sign. In conclusion, we have finally found both x and y:

$$x = \frac{1}{2} \quad \text{and} \quad y = \frac{1}{2}(\phi - 1) = \frac{1}{2}\phi^{-1}$$

(we used the identity $\phi^{-1} = \phi - 1$ to rewrite y).

In conclusion, our rotation, q_5, along the vertex $\mathbf{v}_1$ has the quaternion form

$$\frac{1}{2}\left(\phi + \mathbf{i} + \phi^{-1}\,\mathbf{j}\right).$$

Moreover, since $\mathbf{v}_1$ and $\mathbf{i}$ determine an edge which is parallel to the y-axis we have also found $\mathbf{v}_1$ to be twice the pure part of q_5, i.e.

$$\mathbf{v}_1 = \mathbf{i} + \phi^{-1}\,\mathbf{j}.$$

[III] Note that there was a bit of luck involved in following the particular point p. If I would have followed the pointed centered on the face above it I would end up pushed off of the picture!

[IV] Of course this is why I took the sign of the original 1/2 to be negative: so this the sign of a would end up being positive

Since the rotations $\mathbf{R}_{q_5}$ and $\mathbf{R}_{\mathbf{i}}$ can easily be seen to generate all of $S\left(\text{✶}^3\right)$, and since $\mathbf{i}$ (and also q_5, but we don't care) has even order Notice that together q_5 and $\mathbf{i}$ generate all of $B\left(\text{✶}^3\right)$.[V] That is, in the notation from Chapter 9,

$$B\left(\text{✶}^3\right) = \left\langle \mathbf{i}, \frac{1}{2}\left(\phi + \mathbf{i} + \phi^{-1}\,\mathbf{j}\right)\right\rangle.^{\text{VI}} \tag{12}$$

To calculate the matrix for $\mathbf{R}_{q_5}$ it's useful to notice that since $\phi^2 = \phi + 1$ we have that (dividing both sides by ϕ) $\phi = 1 + \phi^{-1}$ or $\phi - 1 = \phi^{-1}$ and $\phi - \phi^{-1} = 1$ (actually we just used these facts above). Instead dividing both sides by ϕ^{-2} shows us that $1 - \phi^{-1} = \phi^{-2}$. Therefore

$$\mathbf{R}_{q_5}(1,1) = \frac{1}{4}\left(\phi^2 - \phi^{-2} + 1\right) = \frac{1}{4}\left(\phi + 1 - 1 + \phi^{-1} + 1\right) = \frac{1}{4}\left(\phi + \phi^{-1} + 1\right) = \frac{1}{2}\phi$$

$$\mathbf{R}_{q_5}(2,1) = 2\left(\frac{1}{2}\phi^{-1}\cdot\frac{1}{2}\right) = \frac{1}{2}\phi^{-1}$$

$$\mathbf{R}_{q_5}(3,1) = -2\left(\frac{1}{2}\phi\cdot 1\right) = -\frac{1}{2}\phi$$

$$\mathbf{R}_{q_5}(1,2) = 2\left(\frac{1}{2}\phi\cdot\frac{1}{2}\right) = \frac{1}{2}\phi$$

$$\mathbf{R}_{q_5}(2,2) = \frac{1}{4}\left(\phi^2 + \phi^{-2} - 1\right) = \frac{1}{4}\left(\phi + 1 + 1 - \phi^{-1} - 1\right) = \frac{1}{4}\cdot 2 = \frac{1}{2}$$

$$\mathbf{R}_{q_5}(3,2) = 2\left(\frac{1}{2}\phi^{-1}\cdot\frac{1}{2}\right) = \frac{1}{2}\phi^{-1}$$

$$\mathbf{R}_{q_5}(1,3) = -2\left(\frac{1}{2}\phi\cdot\frac{1}{2}\phi^{-1}\right) = -\frac{1}{2}$$

$$\mathbf{R}_{q_5}(2,3) = 2\left(\frac{1}{2}\phi\cdot\frac{1}{2}\phi^{-1}\right) = \frac{1}{2}$$

$$\mathbf{R}_{q_5}(3,3) = \frac{1}{4}\left(\phi^2 - \phi^{-2} - 1\right) = \frac{1}{4}\left(\phi + 1 - 1 + \phi^{-1} - 1\right) = \frac{1}{4}\left(\phi + \phi^{-1} - 1\right) = \frac{1}{2}\phi^{-1}.$$

In conclusion, this means that the matrix is

$$\mathbf{R}_{q_5} = \begin{pmatrix} \frac{\phi}{2} & \frac{\phi^{-1}}{2} & \frac{1}{2} \\ \frac{\phi^{-1}}{2} & \frac{1}{2} & -\frac{\phi}{2} \\ -\frac{1}{2} & \frac{\phi}{2} & \frac{\phi^{-1}}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \phi & \phi^{-1} & 1 \\ \phi^{-1} & 1 & -\phi \\ -1 & \phi & \phi^{-1} \end{pmatrix}$$

Putting these together we see that the symmetries of the dodecahedron (or icosahedron) is[VII]

$$S\left(\text{✶}^3\right) = \left\langle \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} \phi & \phi^{-1} & 1 \\ \phi^{-1} & 1 & -\phi \\ -1 & \phi & \phi^{-1} \end{pmatrix}\right\rangle.$$

[V] If you haven't read Chapter 9, for a subset to generate a group simply means that every element of the group can be written as products of elements in the subset. Also, for this same student, see the exercises for a temporary definition of $B\left(\text{✶}^3\right)$. If you haven't read this chapter, don't panic. We won't need these facts at this time.

[VI] If you haven't read Chapter 9 ignore this for now.

[VII] You won't know what this means until after reading Chapter 9. If you read this before then, just think of this as saying that every symmetry of a dodecahedron is the composition of the two matrices listed.

Let's play with this to figure out what all of the vertices of an icosahedron are. We have already found one vertex, namely $\mathbf{v}_1$: $\mathbf{v}_1 = \mathbf{i} + \phi^{-1}\mathbf{j}$. It will turn out to be slightly more convenient to find the vertices of a slightly dilated icosahedron and we replace

$$\mathbf{v}_1 \rightsquigarrow \phi \cdot \mathbf{v}_1 = \phi\mathbf{i} + \mathbf{j}.$$

This is no big deal since if we don't like this we can simply divide everything by the golden ratio.[VIII] The net effect is to simply uniformly scale the resulting icosahedron by a factor of the golden ratio, ϕ, i.e. the icosahedron is ϕ times larger than the one we would have had if we started with the original $\mathbf{v}_1$.

The next vertex, $\mathbf{v}_2$, which we've said is the opposite vertex on our original edge therefore has coordinates

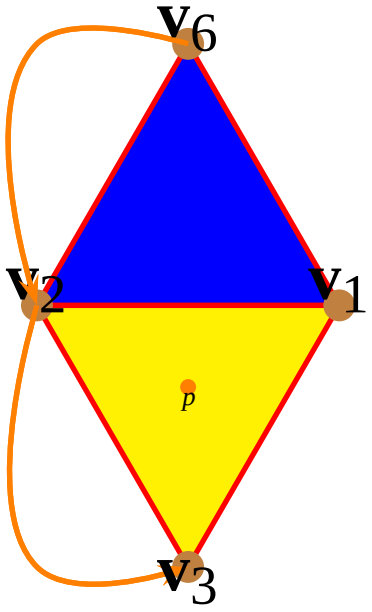
$$\mathbf{v}_2 = \mathbf{i}\mathbf{v}_1\mathbf{i}^{-1} = \phi\,\mathbf{i} - \mathbf{j}.$$

Now we apply combinations of $\mathbf{R}_\mathbf{i}$ and $\mathbf{R}_{q_5}$ to gradually move the vertex $\mathbf{v}_2$ around to be able to find the other $12 - 2 = 10$ vertices of the icosahedron. Doing just that, we search for the other 10 vertices of the icosahedron by applying our symmetries, $\mathbf{R}_{q_5}$ and $\mathbf{R}_\mathbf{i}$, to $\mathbf{v}_1$ and $\mathbf{v}_2$ until we end up with 12 vertices. It's probably faster to make these computations using the matrix. (We write the vertices as column vectors) For instance,

$$\begin{aligned}
\mathbf{R}_{q_5}(\mathbf{v}_2) &= \mathbf{R}_{q_5}\left(\begin{bmatrix} \phi \\ -1 \\ 0 \end{bmatrix}\right) \\
&= \phi\mathbf{R}_{q_5}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) - \mathbf{R}_{q_5}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) \\
&= \frac{\phi}{2} \cdot \begin{bmatrix} \phi \\ \phi^{-1} \\ -1 \end{bmatrix} - \frac{1}{2} \cdot \begin{bmatrix} \phi^{-1} \\ 1 \\ \phi \end{bmatrix} \\
&= \frac{1}{2} \cdot \begin{bmatrix} \phi + 1 \\ 1 \\ -\phi \end{bmatrix} - \frac{1}{2} \cdot \begin{bmatrix} \phi - 1 \\ 1 \\ \phi \end{bmatrix} \\
&= \begin{bmatrix} 1 \\ 0 \\ -\phi \end{bmatrix}
\end{aligned}$$

which we call $\mathbf{v}_3$.

[VIII] Note that $\mathbf{i}$ is no longer the center of an edge, but $\phi\mathbf{i}$ is.



Similarly, we find the other vertices (now that we have the matrices which represent these rotations, it's probably faster to directly use them)!

$$\begin{aligned}
\mathbf{v}_1 &:= \phi\mathbf{i} + \mathbf{j} \\
\mathbf{v}_2 &:= \mathbf{i}\mathbf{v}_1\mathbf{i}^{-1} = \phi\mathbf{i} - \mathbf{j} \\
\mathbf{v}_3 &:= \mathsf{q}_5\mathbf{v}_2\mathsf{q}_5^{-1} = \mathbf{i} - \phi\mathbf{k} \\
\mathbf{v}_4 &:= \mathsf{q}_5\mathbf{v}_3\mathsf{q}_5^{-1} = \phi\mathbf{j} - \mathbf{k} \\
\mathbf{v}_5 &:= \mathsf{q}_5\mathbf{v}_4\mathsf{q}_5^{-1} = \phi\mathbf{j} + \mathbf{k} \\
\mathbf{v}_6 &:= \mathsf{q}_5\mathbf{v}_5\mathsf{q}_5^{-1} = \mathbf{i} + \phi\mathbf{k}
\end{aligned}$$

which can easily be checked by using the matrix for q_5 given above. We could continue this way, but if we notice that the anti-polar map

$$(x, y, z) \mapsto (-x, -y, -z)$$

is a symmetry of the icosahedron, we get a nice shortcut by taking all 6 of the vertices we've found so far and take their negatives to find the other 6 vertices[IX]

$$\begin{aligned}
\mathbf{v}_1' &:= -\mathbf{v}_1 = -\phi\mathbf{i} - \mathbf{j} \\
\mathbf{v}_2' &:= -\mathbf{v}_2 = -\phi\mathbf{i} + \mathbf{j} \\
\mathbf{v}_3' &:= -\mathbf{v}_3 = -\mathbf{i} + \phi\mathbf{k} \\
\mathbf{v}_4' &:= -\mathbf{v}_4 = -\phi\mathbf{j} + \mathbf{k} \\
\mathbf{v}_5' &:= -\mathbf{v}_5 = -\phi\mathbf{j} - \mathbf{k} \\
\mathbf{v}_6' &:= -\mathbf{v}_6 = -\mathbf{i} - \phi\mathbf{k}
\end{aligned}$$

[IX] To save a bit of time, we can see that the negative off all 6 of these vertices are not equal to any of the original vertices since the first component which shows up in any of them is always positive.

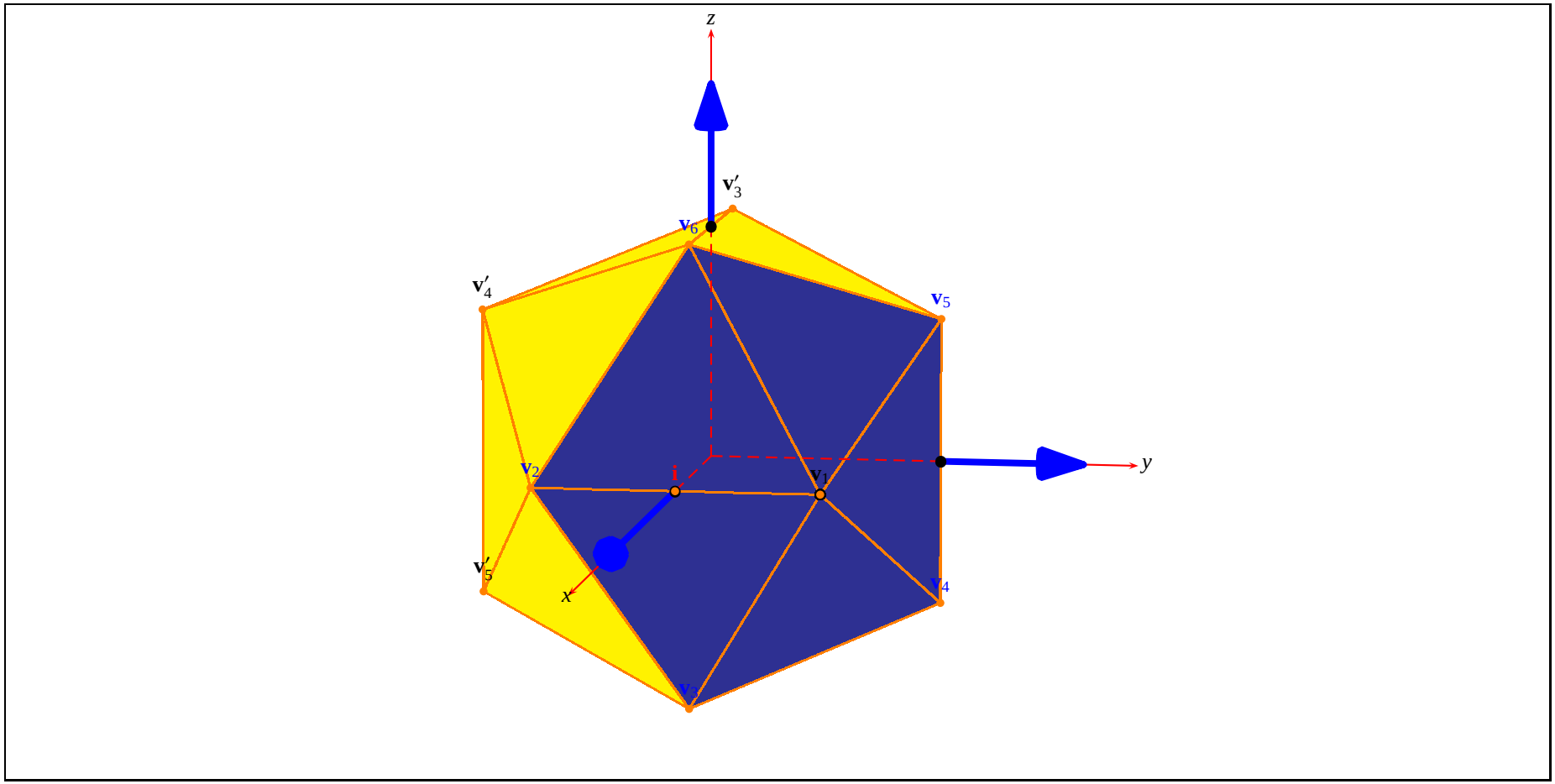


Figure C.6: A labeled icosahedron. Note that the 5 blue vertices, $\mathbf{v}_i$ ($i = 2, 3, 4, 5, 6$) surround $\mathbf{v}_1$ in a pentagon shape going clockwise. This matches the fact that these vertices were found by using $\mathfrak{q}_5$, which is a $\frac{2}{5}\pi$ rotation about the vertex $\mathbf{v}_1$.

are also vertices of the icosahedron. Observe that

$$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6, \mathbf{v}_1', \mathbf{v}_2', \mathbf{v}_3', \mathbf{v}_4', \mathbf{v}_5', \mathbf{v}_6' \tag{13}$$

make up 12 distinct vertices of the icosahedron, and hence, since an icosahedron has 12 vertices, are all of the vertices of the icosahedron.

REMARK C.46.

(1) We can consciously write the set of vertices as $(\pm\phi, \pm 1, 0)$, $(0, \pm\phi, \pm 1)$, and $(\pm 1, 0, \pm\phi)$. Notice that direction which has a ϕ of it is always cyclically before the term which does not! For instance, $(1, \phi, 0)$ is not a vertex of $\star^3$.

(2) Notice the dot product of vertices. $\mathbf{v}$ and $\mathbf{w}$ can take on one of the following values:

- $\mathbf{v} \bullet \mathbf{w} = \phi^2 + 1 = \phi + 2 \Leftrightarrow \mathbf{v} = \mathbf{w}$,
- $\mathbf{v} \bullet \mathbf{w} = \phi \Leftrightarrow \mathbf{v}$ and $\mathbf{w}$ are connected by an edge,
- $\mathbf{v} \bullet \mathbf{w} = -\phi \Leftrightarrow \mathbf{v}$ and $-\mathbf{w}$ are connected by an edge,
- $\mathbf{v} \bullet \mathbf{w} = -(\phi^2 + 1) = -\phi - 2 \Leftrightarrow \mathbf{v} = -\mathbf{w}$.

Note also that a triple of vertices form the vertices of a face of the icosahedron iff all dot products between are equal to ϕ. For convenience I record the dot products of vertices in the following 12 by 12 table whose i j entry is $\mathbf{v}_i \bullet \mathbf{v}_j$: (Highlighted in the table in yellow the vertices $\mathbf{v}_2$, $\mathbf{v}_6$ and $\mathbf{v}'_4$ are the vertices of a face since each dot product of distinct pairs there is ϕ.)

$\bullet$	$\mathbf{v}_1$	$\mathbf{v}_2$	$\mathbf{v}_3$	$\mathbf{v}_4$	$\mathbf{v}_5$	$\mathbf{v}_6$	$\mathbf{v}'_1$	$\mathbf{v}'_2$	$\mathbf{v}'_3$	$\mathbf{v}'_4$	$\mathbf{v}'_5$	$\mathbf{v}'_6$
$\mathbf{v}_1$	$2+\phi$	ϕ	ϕ	ϕ	ϕ	ϕ	$-2-\phi$	$-\phi$	$-\phi$	$-\phi$	$-\phi$	$-\phi$
$\mathbf{v}_2$	ϕ	$2+\phi$	ϕ	$-\phi$	$-\phi$	ϕ	$-\phi$	$-2-\phi$	$-\phi$	ϕ	ϕ	$-\phi$
$\mathbf{v}_3$	ϕ	ϕ	$2+\phi$	ϕ	$-\phi$	$-\phi$	$-\phi$	$-\phi$	$-2-\phi$	$-\phi$	ϕ	ϕ
$\mathbf{v}_4$	ϕ	$-\phi$	ϕ	$2+\phi$	ϕ	$-\phi$	$-\phi$	ϕ	$-\phi$	$-2-\phi$	$-\phi$	ϕ
$\mathbf{v}_5$	ϕ	$-\phi$	$-\phi$	ϕ	$2+\phi$	ϕ	$-\phi$	ϕ	ϕ	$-\phi$	$-2-\phi$	$-\phi$
$\mathbf{v}_6$	ϕ	ϕ	$-\phi$	$-\phi$	ϕ	$2+\phi$	$-\phi$	$-\phi$	ϕ	ϕ	$-\phi$	$-2-\phi$
$\mathbf{v}'_1$	$-2-\phi$	$-\phi$	$-\phi$	$-\phi$	$-\phi$	$-\phi$	$2+\phi$	ϕ	ϕ	ϕ	ϕ	ϕ
$\mathbf{v}'_2$	$-\phi$	$-2-\phi$	$-\phi$	ϕ	ϕ	$-\phi$	ϕ	$2+\phi$	ϕ	$-\phi$	$-\phi$	ϕ
$\mathbf{v}'_3$	$-\phi$	$-\phi$	$-2-\phi$	$-\phi$	ϕ	ϕ	ϕ	ϕ	$2+\phi$	ϕ	$-\phi$	$-\phi$
$\mathbf{v}'_4$	$-\phi$	ϕ	$-\phi$	$-2-\phi$	$-\phi$	ϕ	ϕ	$-\phi$	ϕ	$2+\phi$	ϕ	$-\phi$
$\mathbf{v}'_5$	$-\phi$	ϕ	ϕ	$-\phi$	$-2-\phi$	$-\phi$	ϕ	$-\phi$	$-\phi$	ϕ	$2+\phi$	ϕ
$\mathbf{v}'_6$	$-\phi$	$-\phi$	ϕ	ϕ	$-\phi$	$-2-\phi$	ϕ	ϕ	$-\phi$	$-\phi$	ϕ	$2+\phi$

(3) In particular, no two vertices (have vectors which) are orthogonal. This means that the vertices of an icosahedron do not contain the vertices of any octahedron. This is despite that the dual of an icosahedron, the dodecahedron, contains vertices which form the vertices of the dual of an icosahedron, the cube.

EXERCISE C.14. Write the (unit) quaternion

$$\frac{1}{6} + \frac{5}{6}\mathbf{i} + \frac{1}{2}\mathbf{j} + \frac{1}{6}\mathbf{k}$$

as a rotation matrix.

SOLUTION

Exercise C.15. Describe the rotations $\mathbf{R}_\mathbf{i}$, $\mathbf{R}_\mathbf{j}$, and $\mathbf{R}_\mathbf{k}$. Prove that $\{\mathbf{1}_{\mathbb{R}^3}, \mathbf{R}_\mathbf{i}, \mathbf{R}_\mathbf{j}, \mathbf{R}_\mathbf{k}\}$ is a group (of order 4). Describe an object which has this group as a symmetry group.
Hint! It could be faster to work backwards!

SOLUTION

Exercise C.16. Use quaternions to find the axis of rotation of a 45 degree (counter-clockwise) rotation along the x-axis followed by a 60 degree (counter-clockwise) rotation along the y-axis.

SOLUTION

Exercise C.17. [x] Define $B\left(\text{✶}^3\right)$ as the set of quaternions q which induce symmetries of an icosahedron. (This is the same as requiring $\mathbf{R}_\mathsf{q}$ permutes the vertices of the icosahedron, (13), since a rotation – which $\mathbf{R}_\mathsf{q}$ is an example of – is a symmetry of a regular polyhedron iff it permutes its vertices.)

SOLUTION

(1) Prove that $B\left(\text{✶}^3\right)$ is a group.

(2) Prove that $B\left(\text{✶}^3\right)$ has 120 elements.

(3) Show none of the quaternions

$$\frac{1}{\sqrt{2}}(1+\mathbf{i}) \qquad \frac{1}{\sqrt{2}}(1+\mathbf{j}) \qquad \frac{1}{\sqrt{2}}(1+\mathbf{k})$$

are an element of $B\left(\text{✶}^3\right)$.

(4) Show that

$$\omega = \frac{1}{2} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{i} + \frac{1}{2}\mathbf{k}$$

is an element of $B\left(\text{✶}^3\right)$. (Just show that $\mathbf{R}_\omega$ permutes the vertices of the icosahedron.)

Since $\mathbf{R}_\omega$ has order 3, we can take the (a scaled version of) pure part of ω, $\mathbf{i}+\mathbf{j}+\mathbf{k} = \begin{bmatrix}1\\1\\1\end{bmatrix}$ as the center of some face of an icosahedron, and since the dodecahedron is dual to the icosahedron, the vector is a vertex of a dodecahedron.

Exercise C.18.

SOLUTION

(1) Find the other vertices of the dodecahedron.

[x] You need to have covered Chapter 5, Group Basics, to be able to do this problem. The group $B\left(\text{✶}^3\right)$ is first introduced in the text in Definition 8.21 on page 387.

(2) Prove that the dodecahedron contains 5 cubes whose vertices are vertices of the dodecahedron. Prove that these five cubes are permuted by the symmetries of the dodecahedron.
Hint! It might be easier to work backwards. If you got the correct vertices, there is an obvious cube. Then use symmetries to push around the original cube to get 4 others.

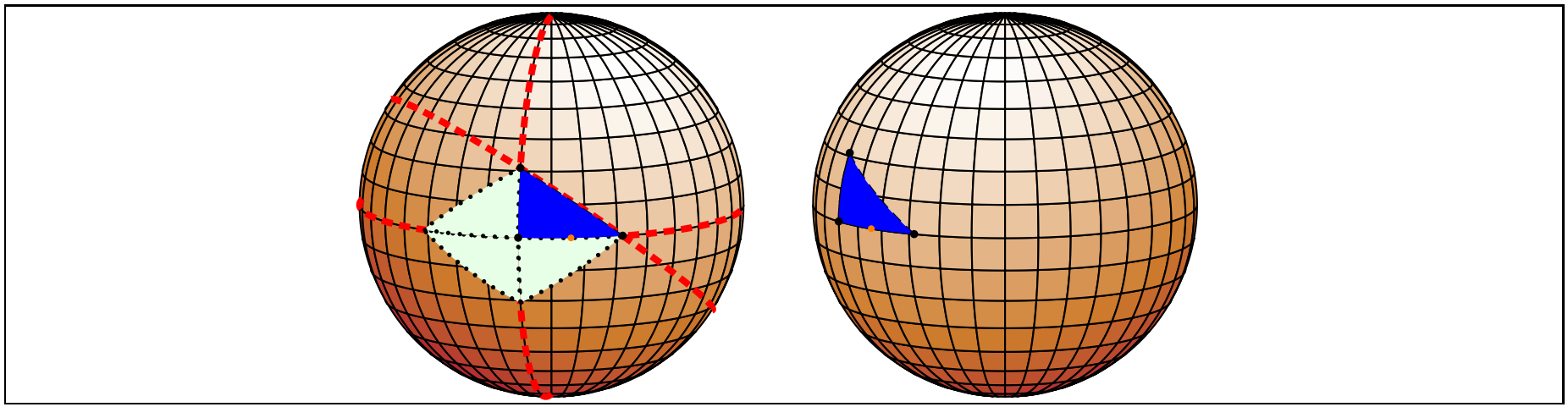


Figure C.7: The planes of reflection which determine $S\left(\text{☆}^3\right)$ and $S^*\left(\text{☆}^3\right)$ meet at the indicated angles. You can think of this triangle as being a spherical triangle, i.e. a triangle that lives on the surface of a ball (i.e. the Earth). Starting with the orange point and connecting the orbit of that point under the action of $S^*\left(\text{☆}^3\right)$ gives rise to various polyhedra. For instance, if we start at the point in the bottom right corner, we end up with an icosahedron. If we start with the point the the upper left hand corner, you end up with a dodecahedron. The exercise below, Exercise C.19, asks, in a sense, what we get when we start with the indicated point.

Many interesting polyhedrons are found by taking a point $\mathbf{v} \in \mathbb{R}^3$ and considering the polyhedron with vertices

$$S\left(\text{☆}^3\right) \bullet \mathbf{v} \stackrel{\text{def}}{=} \left\{\rho(\mathbf{v}) : \rho \in S\left(\text{☆}^3\right)\right\}.$$

Note this is the same thing as

$$\left\{\mathbf{R}_{\mathsf{q}}(\mathbf{v}) : \mathsf{q} \in B\left(\text{☆}^3\right)\right\}.$$

In the language of Chapter 10 this is the orbit of the point $\mathbf{v}$ under the action of $S\left(\text{☆}^3\right)$ (or $B\left(\text{☆}^3\right)$). We have already seen the icosahedron and the dodecahedron in this way: as $S\left(\text{☆}^3\right) \bullet \begin{bmatrix} \phi \\ 1 \\ 0 \end{bmatrix}$ and $S\left(\text{☆}^3\right) \bullet \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ respectively.

In fact all reasonable polyhedron are obtained in this way by using the three types of $\mathbb{R}^3$ symmetry.

EXERCISE C.19. Describe

$$S\left(\text{☆}^3\right) \bullet \begin{bmatrix} 3\phi \\ 1 \\ 0 \end{bmatrix}.$$

SOLUTION

Hint! Your ¡gooaaaaaal! is to understand this subtle hint.
Hint! Attempt to draw this inside of a dodecahedron by taking a point which is halfway between the center of a face and the center of an edge and passing it around the entire dodecahedron by applying symmetries.

D Regular Polyhedra and Platonic Solids

There are a total of 5 regular platonic solids: the tetrahedron– denoted $\blacktriangleright^3$, the cube – denoted $\blacklozenge^3$ – (i.e. the hexahedron), the octahedron – denoteed $\blacklozenge^3$, the icosahedron– denoted $\bigstar^3 4$, and the dodecahedron – denoted $\bigstar^3$. While these shapes are, of course, more complicated then the regular polygons that we are familar with from 2 dimensional geometry, it's rather remarkable that there are only 5 regular polyhedra in 3 dimensions, but there are an infinite number of them in 2 dimensions.[I]

Name	Symbol	Vertices	Edges	Faces	$\lvert S^* \rvert$	$\lvert S \rvert$
Tetrahedron	$\blacktriangleright^3$	4	6	6	24	12
Octahedron	$\blacklozenge^3$	6	12	8	48	24
Cube	$\blacklozenge^3$	8	12	6[II]	48	24
Icosahedron	$\bigstar^3$	12	30	20	120	60
Dodecahedron	$\bigstar^3$	20	30	12	120	60

Table D.1: Basic Data for Platonic Solids

D.1 The Cube

D.1.1 Description

The (standard unit) *square*, denoted $\blacklozenge^2$, can be described as the subset of $\mathbb{R}^2$ as

$$\{(x,y) : |x|, |y| \leq 1\},$$

and the (standard unit) *cube*, denoted $\blacklozenge^3$, can be described as the subset of $\mathbb{R}^3$ as

$$\{(x,y,z) : |x|, |y|, |z| \leq 1\}.$$

[I] It turns out there are 6 in 4 dimensions, then for every dimension after 4 there are are only three!!: the higher dimensional analogue of a triangle, of a cube, and of a diamond. To me, this is one of the most shocking facts in all of mathematics!

[II] Think of the 6 sides of a die

In 3 dimensions, it has 8 vertices

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} \quad \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \quad \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}. \tag{2}$$

More generally, the (standard unit) *n-cube*, denoted $\blacklozenge^n$, can be described as the subset of $\mathbb{R}^n$ given by

$$\{(x_1, \ldots, x_n) : |x_i| \leq 1,\ 1 \leq i \leq n\}.$$

D.1.2 Symmetries

For each of the platonic solids (or any other highly symmetric solid) there are two ways we can count the number of symmetries

(1) From Theorem C.18 back on page 772 in Appendix C, we know that every orientation preserving symmetry of a platonic solid is a rotation. Therefore, to count the number of symmetries an object can have we simply need to count the possible number of rotation axis.

(2) The easier way of doing is to use the **OST**, Theorem 10.94 back on page 554, from Chapter 10. Simply take

$$\begin{aligned} |S(X)| &= \#\text{Vertices of } X \cdot \#\text{Vertices adjacent to a given vertex} \\ \text{or } |S(X)| &= \#\text{Edges of } X \cdot \#\text{Vertices on a given edge} \\ \text{or } |S(X)| &= \#\text{Faces of } X \cdot \#\text{Vertices on a given face.} \end{aligned}$$

Of course the middle equation is the easiest since there are always two vertices per edge and so we have

$$|S(X)| = \#\text{Edges of } X \cdot \#\text{Vertices on a given edge} = 2 \cdot \#\text{Edges of } X.$$

We use both methods for demonstrative purposes:

(1) A cube has 8 vertices which come in 4 anti-polar pairs. About each of these anti-polar pairs

Starting with Chapter 5, this problem can be done at any time, but the longer one waits the easier it will become.

Exercise D.1. Prove that

$$\left|S^*(\blacklozenge^n)\right| = 2^n \cdot n! \quad \text{and} \quad \left|S(\blacklozenge^n)\right| = 2^{n-1} \cdot n!.$$

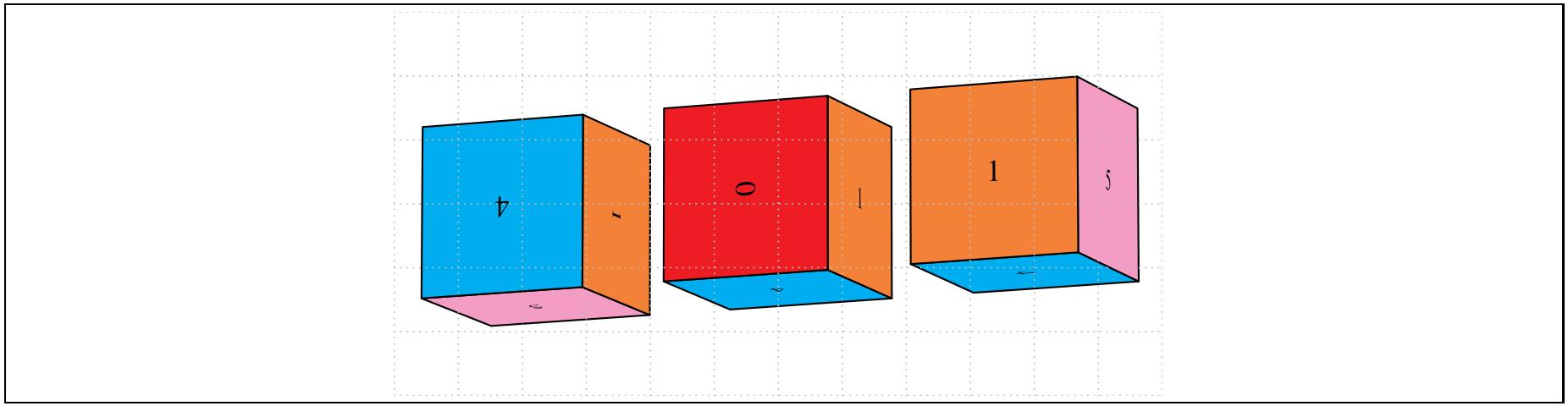


Figure D.1: Rotating a Cube. This is the same figure as Figure C.4

D.2 The Tetrahedron

D.2.1 Description

Just like a cube/square, the (equilateral) triangle/tetrahedron can be generalized to any dimension. In $\mathbb{R}^n$, it is defined as a polyhedron with $n+1$ vertices with any pair of which are the same distance apart. For example, the (equilaterial) triangle $\blacktriangleright^2$ has vertices given by the three cubed roots of unity which make up $\{\{n\}\}_3$:

$$1, \qquad -\frac{1}{2}+\frac{\sqrt{3}}{2}i, \qquad -\frac{1}{2}-\frac{\sqrt{3}}{2}i.$$

In three dimensions we get lucky because we can find a nice tetrahedron whose vertices are all at integer coefficients (so there is no need for this ugly square root which shows up here!) In higher dimensions it's not quite as clear there is even such a thing as such an equilateral object (there is, see the rant below!)

Rant D.1. The generalization of a triangle in 2–dimensions, and a tetrahedron in 3-dimensions is called an n-simplex, τ^n. Perhaps the vertices of $\blacktriangleright^n$ are best represented by going up one dimension to $\mathbb{R}^{n+1}$. They are given by the $n+1$ standard basis vectors:

$$\mathbf{e}_1,\ \mathbf{e}_2,\ \ldots\ ,\ \mathbf{e}_n,\ \mathbf{e}_{n+1}.$$

Note there in this situation we have that any pair of distinct vertices lie at a distance of $\sqrt{2}$ from each other.

Thinking this way, the triangle (in $\mathbb{R}^3$) can be seen as

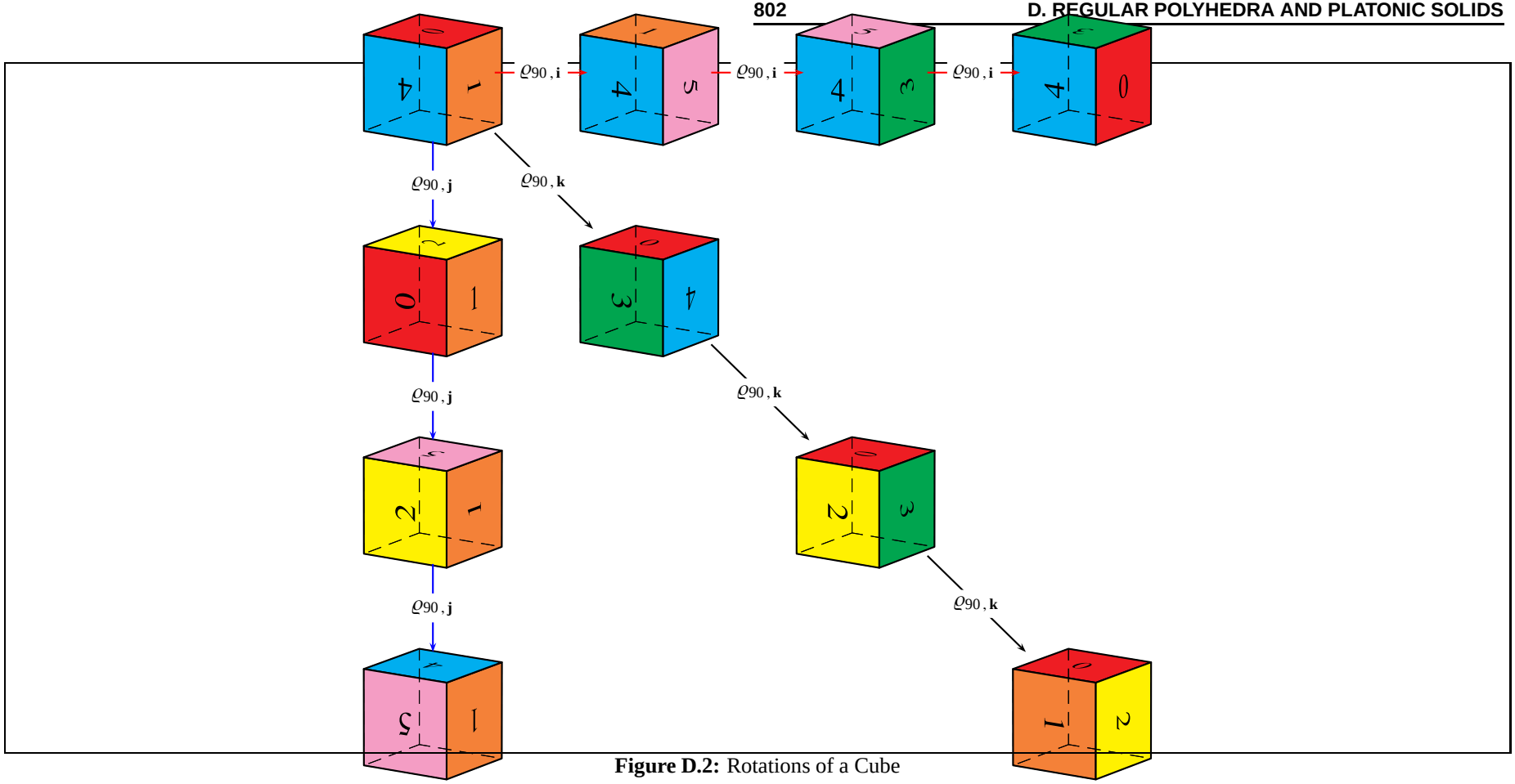
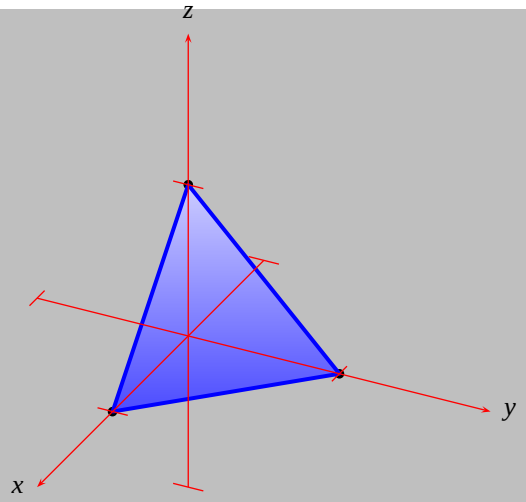


Figure D.2: Rotations of a Cube



The advantage of thinking this way is that it is now clear that there are a total of $n!$ symmetries of $\blacktriangleright^{n-1}$ and $n!/2$ orientation preserving ones. In fact the (total) symmetry group of $\blacktriangleright^{n-1}$ is isomorphic to $\mathcal{S}_n$ (identifying $\mathcal{S}_n$ with permutation matrices in $\mathbf{O}_n(\mathbb{R})$). This generalizes the fact that $\mathcal{D}_3 \cong \mathcal{S}_3$. Indeed, any permutation matrix, by definition, permutes the standard basis vectors $\mathbf{e}_i$ which are what we took to be the vertices of $\blacktriangleright^{n-1}$. On the other hand every symmetry of $\blacktriangleright^{n-1}$ (with vertices taken as above) must permute its vertices, i.e. the standard basis vectors. Meanwhile the orientation preserving symmetries, i.e. those with positive determinant, correspond exactly with the set of even permutations (recall that in Chapter 8 we defined the sign of a permutation as the determinant of the corresponding permutation matrix! Therefore, we have proven:

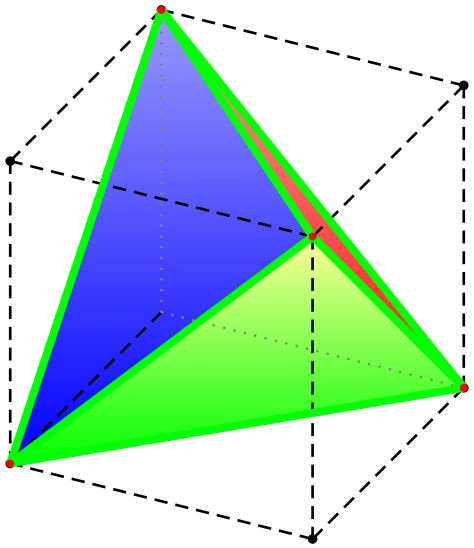
THEOREM D.2. $S^*(\blacktriangleright^{n-1}) \cong \mathcal{S}_n$ *and* $S(\blacktriangleright^{n-1}) \cong \mathcal{A}_n$. ■

Granted these two cool advantages, (identifying the vertices and group of symmetries quickly) there's an obvious disadvantage, namely that we have to go up one dimension. Indeed, in this way to see even the tetrahedron we need to go up to 4 dimensions, which gives most of us a decent amount of trouble!

Let's get back to discussing what we fix as the standard tetrahedron, denoted $\blacktriangleright^3$. It is the polyhedron with four vertices:

$$\begin{bmatrix}1\\1\\1\end{bmatrix} \quad \begin{bmatrix}1\\-1\\-1\end{bmatrix} \quad \begin{bmatrix}-1\\1\\-1\end{bmatrix} \quad \begin{bmatrix}-1\\-1\\1\end{bmatrix}.$$

Note that this works since any pair of distinct vertices are a distance $2\sqrt{2}$ apart from each other. Colloquially, it is a pyramid with a triangular base. Note that its vertices are subset of the vertices of the cube, i.e. the tetrahedron can be inscribed inside of the cube. This is pictured as follows:



D.2.2 Symmetries

For our standard tetrahedra, we have that both $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\begin{bmatrix}1\\-1\\-1\end{bmatrix}$ are vertices which are adjacent (in the tetrahedron, any two vertices are adjacent). Therefore their midpoint,

$$\frac{1}{2}\left(\begin{bmatrix}1\\1\\1\end{bmatrix}+\begin{bmatrix}1\\-1\\-1\end{bmatrix}\right)=\mathbf{i}$$

is the center of a vertex. Therefore, $B\left(\blacktriangleright^3\right)^3$ admits a symmetry which is 180° rotation about $\mathbf{i}$, i.e. it's just the quaternion

$$\cos\pi/2+\sin\pi/2\mathbf{i}=\mathbf{i}\in B\left(\blacktriangleright^3\right).$$

Moreover, ever vertex has degree 3, so rotation by $2\pi/3 = 120° = \frac{360°}{3}$ degrees along the axis connecting the origin to any of these vertices is a symmetry of $B\left(\blacktriangleright^3\right)^3$.

Therefore, picking the vertex $\begin{bmatrix}1\\1\\1\end{bmatrix}$ we see that

$$\cos\left(\frac{1}{2}\cdot\frac{2\pi}{3}\right)+\sin\left(\frac{1}{2}\cdot\frac{2\pi}{3}\right)\frac{(\mathbf{i}+\mathbf{j}+\mathbf{k})}{\|(\mathbf{i}+\mathbf{j}+\mathbf{k})\|}=\frac{1}{2}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{i}+\frac{1}{2}\mathbf{k}\in B\left(\blacktriangleright^3\right).$$

D.3 The Octahedron

D.3.1 Description

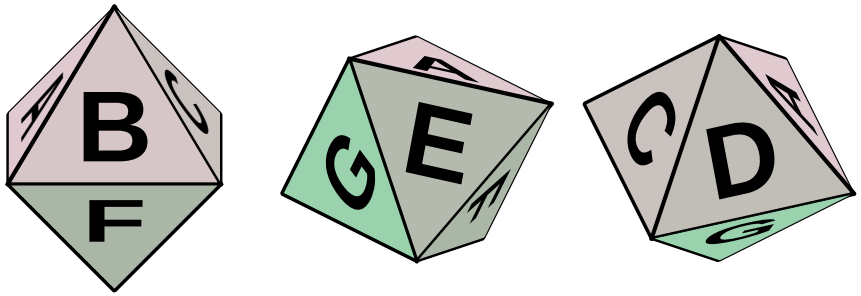
The Octahedron, denoted $\blacklozenge^3$, is the dual polyhedron of the cube. It has six vertices:

$$\begin{bmatrix}1\\0\\0\end{bmatrix}\quad\begin{bmatrix}-1\\0\\0\end{bmatrix}\quad\begin{bmatrix}0\\1\\0\end{bmatrix}\quad\begin{bmatrix}0\\-1\\0\end{bmatrix}\quad\begin{bmatrix}0\\0\\1\end{bmatrix}\quad\begin{bmatrix}0\\0\\-1\end{bmatrix}.$$

This polyhedra has an obvious generalization to higher dimensions, it's simply the polytope with $2n$ vertices given by

$$\pm\mathbf{e}_1,\ldots,\pm\mathbf{e}_n.$$

Colloquially, it consists of two pyramids with square bases glued together along those bases.



D.3.2 Symmetries

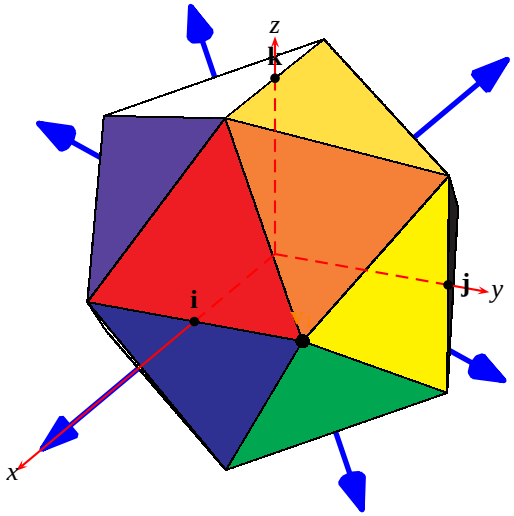
D.4 The Icosahedron

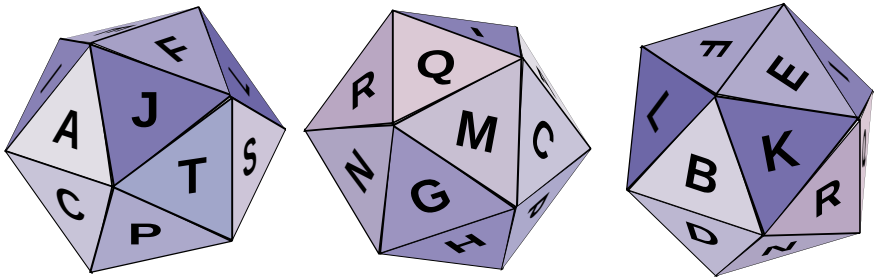
D.4.1 Description

On page 792 we showed that if we line up an icosahedron in a fairly obvious way, then it has vertices:

$$\begin{bmatrix}\phi\\1\\0\end{bmatrix}\quad\begin{bmatrix}\phi\\-1\\0\end{bmatrix}\quad\begin{bmatrix}\phi\\-1\\0\end{bmatrix}\quad\begin{bmatrix}-\phi\\-1\\0\end{bmatrix}$$
$$\begin{bmatrix}0\\\phi\\1\end{bmatrix}\quad\begin{bmatrix}0\\\phi\\-1\end{bmatrix}\quad\begin{bmatrix}0\\\phi\\-1\end{bmatrix}\quad\begin{bmatrix}0\\-\phi\\-1\end{bmatrix}$$
$$\begin{bmatrix}1\\0\\\phi\end{bmatrix}\quad\begin{bmatrix}-1\\0\\\phi\end{bmatrix}\quad\begin{bmatrix}1\\0\\-\phi\end{bmatrix}\quad\begin{bmatrix}-1\\0\\-\phi\end{bmatrix}$$

It is denoted .



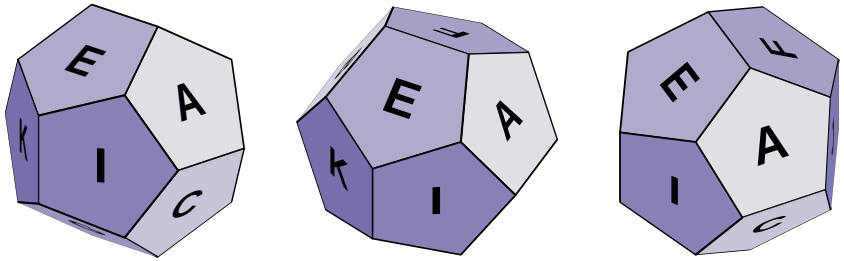


D.4.2 Symmetries

D.5 The Dodecahedron

D.5.1 Description

The dodecahedron is the dual of the icosahedron. It is denoted as $\{5,3\}$. It has 20 vertices and 12 faces.



In Exercise C.18 back on page 795 you should have shown that the twenty vertices of dodecahedron are

$$\mathbf{V}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{V}_2 = \begin{bmatrix} \phi \\ 0 \\ \phi^{-1} \end{bmatrix} \quad \mathbf{V}_3 = \begin{bmatrix} \phi \\ 0 \\ -\phi^{-1} \end{bmatrix} \quad \mathbf{V}_4 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \quad \mathbf{V}_5 = \begin{bmatrix} \phi^{-1} \\ \phi \\ 0 \end{bmatrix}$$

$$\mathbf{V}_6 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \quad \mathbf{V}_7 = \begin{bmatrix} 0 \\ \phi^{-1} \\ -\phi \end{bmatrix} \quad \mathbf{V}_8 = \begin{bmatrix} -\phi^{-1} \\ \phi^{-1} \\ 0 \end{bmatrix} \quad \mathbf{V}_9 = \begin{bmatrix} 0 \\ \phi^{-1} \\ \phi \end{bmatrix} \quad \mathbf{V}_{10} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$\mathbf{V}_{11} = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} \quad \mathbf{V}_{12} = \begin{bmatrix} -\phi \\ 0 \\ -\phi^{-1} \end{bmatrix} \quad \mathbf{V}_{13} = \begin{bmatrix} -\phi \\ 0 \\ \phi^{-1} \end{bmatrix} \quad \mathbf{V}_{14} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \quad \mathbf{V}_{15} = \begin{bmatrix} -\phi^{-1} \\ -\phi \\ 0 \end{bmatrix}$$

$$\mathbf{V}_{16} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{V}_{17} = \begin{bmatrix} 0 \\ -\phi^{-1} \\ \phi \end{bmatrix} \quad \mathbf{V}_{18} = \begin{bmatrix} \phi^{-1} \\ -\phi^{-1} \\ 0 \end{bmatrix} \quad \mathbf{V}_{19} = \begin{bmatrix} 0 \\ -\phi^{-1} \\ -\phi \end{bmatrix} \quad \mathbf{V}_{20} = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}.$$

D.5.2 Symmetries

E Rings

Most students find rings to be slightly easier than groups. This is dispite the fact that rings have two operations and groups have only 1. The reason for this is probably linked to the fact that you guys are probably familiar with many more examples of "real rings." For instance, as a group the integers are cyclic. This is the most basic kind of group. To get an insight into what kinds of things can happen in a general group the integers suck as a test case. The best test case for a group is probably the symmetric group. In fact the entire theory of groups sprung from the study of $\mathcal{S}_n$. On the other hand the integers are a very interesting ring. For starters this ring does not have zero-divisors, but only two of its elements are invertible. In fact much of the theory of rings can be thought of as a generalization of what happens in the integers (and polynomial rings). For instance in ring theory one often studies prime elements and factorizations: two concepts whose definition is identical to what one finds in the integers.

Since most students have little to no background coming into an abstract algebra course with $\mathcal{S}_n$, it's difficult to motivate many facets of its generalization, group theory. However it is much easier to motivate many facets of ring theory. In several recent abstract algebra textbooks ring theory is actually covered *before* group theory. I personally think this a great idea. This is despite the fact that I consider myself a group theorist and "like" groups better than rings. However it is difficult to break this trend in since professors rarely teach both a student's first and second course in abstract algebra. Therefore unless there is a standard agreement, the poor sap of a student might end up learning either group theory or ring theory twice!

E.1 Definitions

A ring is a triple $(R, +, \cdot)$ where $+$ and $\cdot$ are two binary operations

$$\cdot : R \times R \to R$$
$$+ : R \times R \to R$$

which are generically called multiplication and addition, respectively. These operations must satisfy

(1) $(R, +)$ is a group.

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

(3) The distributive properties holds: for all $a, b, r \in R$,

$$(a + b) \cdot r = a \cdot r + b \cdot r$$
$$r \cdot (a + b) = r \cdot a + r \cdot b.$$

For a ring R, the identity of the monoid $(R, \cdot)$ (resp. the group $(R, +)$) is denoted as 1_R (resp. 0_R). It is called the multiplicative (resp. additive) identity.
Let's expand the definition of a ring as much as possible.

DEFINITION E.1. (**Expanded Definition of a Ring**) Let R be a set with two binary operations $+$ and $\cdot$. We say that $(R, +, \cdot)$ is a *ring* if for all $a, b, c \in R$

Associative
- $(a + b) + c = a + (b + c)$
- $a \cdot (b \cdot c) = (a \cdot b) \cdot c.$

Identities
- $\exists 0_R \in R$ s.t. , $a + 0_R = 0_R + a = a$
- $\exists 1_R \in R$ s.t. $a = a \cdot 1_R = 1_R \cdot a.$

Inverses
- $\exists s \in S$ s.t. $a + s = s + a = 0_R.$

Distributive
- $(a + b) \cdot c = a \cdot c + b \cdot c$ $a \cdot (b + c) = a \cdot b + a \cdot c.$

EXERCISE E.1. Let X be a non-empty set.

SOLUTION

(1) Explain why neither $(\mathcal{P}(X), \cup, \cap)$ or $(\mathcal{P}(X), \cap, \cup)$ are rings.

(2) For the two examples from the last part, which of the defining properties hold? Which do not hold?

REMARK E.2. Whenever talking about group $(R, +)$ we use additive notation. Therefore we write $2r$ as $r + r$ and r^2 means $r \cdot r$. As such, the inverse of an element r is denoted as $-r$, etc. Notice with the notation established $1_R \cdot r$ and $1r$ are two different ways of writing the same element, namely r.

The glaring asymmetry in the definition is that not every element is required to have a multiplicative inverse. The set of elements which do have a multiplicative inverse are called *units*. As we have been doing, this is denoted as $\mathcal{U}(R)$.

THEOREM E.3. *For any ring R, $\mathcal{U}(R)$ forms a group.*

PROOF. This is the set of invertible elements in the monoid $(R, \cdot)$, therefore we can apply Corollary 4.39 back on page 228 to get the result. ■

WARNING! E.4. For a ring R, $\mathcal{U}(R)$ will always correspond to the elements with a multiplicative inverse and never to the elements with an additive inverse. Indeed, this latter case would be silly since by definition of a ring, every element has an additive inverse.

As you should be able to predict a subring of a ring is a ring inside of a ring. The only problem is what we should do with the identity, 1_R. We require that the identity of R belongs to any subring of R. Luckily, we won't be discussing rings long enough for us to annoy us either way (in a longer course on ring theory typically this is defined initially without this extra hypothesis and as the course progresses there is a gradual shift to the definition presented here).

DEFINITION E.5. A subset S of a ring R is a *subring* if

(1) $1_R \in S$

(2) S forms a ring under (the restriction of) the operations from R.

NOTATION E.6. The notation $S \preccurlyeq R$ means that S is a subring of R.

As you would hope, we have a subring recognition test:

THEOREM E.7. *$S \preccurlyeq R$ iff S is closed under multiplication, subtraction, and contains 1_R.* ■

EXAMPLE E.8.

(1) For each positive integer n, $\mathbb{Z}_n$ is a ring under (modular) addition and multiplication.

(2) We have $\mathbb{Z} \preccurlyeq \mathbb{Q} \preccurlyeq \mathbb{R} \preccurlyeq \mathbb{C} \preccurlyeq \mathbb{H}$.

(3) If R is a ring, then so is $\mathbf{M}_n(R)$, the set of n by n matrices with coefficients in R.

(4) The set of functions from $\mathbb{R} \to \mathbb{R}$, $\mathcal{F}_{\mathbb{R}}$, forms a ring under addition and multiplication. However, it does not form a ring under addition and composition. The only property which fails to hold is the right distributive property. For instance if we let

$$\begin{aligned} f &: \mathbb{R} \to \mathbb{R} \\ &: x \mapsto x^2 \end{aligned} \qquad \begin{aligned} g &: \mathbb{R} \to \mathbb{R} \\ &: x \mapsto x^2 \end{aligned} \qquad \begin{aligned} h &: \mathbb{R} \to \mathbb{R} \\ &: x \mapsto x \end{aligned}$$

then

$$\begin{aligned} (f \circ (g+h))(x) &= f((g+h)(x)) = f(x^2+x) = (x^2+x)^2 \qquad \text{but} \\ (f \circ g + f \circ h)(x) &= f(g(x)) + f(h(x)) = f(x^2) + f(x) = x^4 + x. \end{aligned}$$

and the two functions are not equal.

(5) The set

$$\mathbb{Z}\left[\frac{1}{2}\right] \stackrel{\text{def}}{=} \left\{\frac{a}{2^n} : a \in \mathbb{Z}\right\}$$

is a ring (and a subring of $\mathbb{Q}$). Indeed, suppose $\frac{a}{2^n}$ and $\frac{b}{2^m}$ are arbitrary element of this set. WLOG, assume that $n \geq m$. Then

$$\begin{aligned}\frac{a}{2^n} + \frac{b}{2^m} &= \frac{a}{2^n} + \frac{b2^{n-m}}{2^n} \\ &= \frac{a}{2^n} + \frac{b2^{n-m}}{2^n} \\ &= \frac{a + b2^{n-m}}{2^n} \\ &= \frac{c}{2^n}\end{aligned}$$

where $c = a + b \cdot 2^{n-m} \in \mathbb{Z}$ (here, we're using $n \geq m$), and $\mathbb{Z}[\frac{1}{2}]$ is closed under addition.

Proving closure under multiplication (for once!!) is easier.

The following theorem justifies the use of additive notation in the definition of a ring.[I]

PROPOSITION E.9. *For any ring, the group $(R, +)$ is abelian.*

In other words for any $a, b \in R$ we have that $a + b = b + a$.

PROOF. Let $a, b \in R$. We need to show that $a + b = b + a$.

$$\begin{aligned}(a + b) \cdot (1_R + 1_R) &= (a + b) \cdot 1_R + (a + b) \cdot 1_R && \text{(Lt. dist. prop.)} \\ &= (a \cdot 1_R) + (b \cdot 1_R) + (a \cdot 1_R) + (b \cdot 1_R) && \text{(Rt distr. prop.)} \\ &= a + b + a + b. && (1_R \text{ is the identity})\end{aligned}$$

Now we do the same calculation but this time we first use the right distributive property and then the left one:

$$\begin{aligned}(a + b) \cdot (1_R + 1_R) &= a \cdot (1_R + 1_R) + b \cdot (1_R + 1_R) && \text{(Rt. dist. prop.)} \\ &= (a \cdot 1_R) + (a \cdot 1_R) + (b \cdot 1_R) + (b \cdot 1_R) && \text{(Lt dist. prop.)} \\ &= a + a + b + b && (1_R \text{ is the identity.})\end{aligned}$$

Comparing both equations we see that

$$a + b + a + b = a + a + b + b.$$

If we subtract a on the left of both sides then subtract b on the right of both sides we finially see that

$$b + a = a + b$$

as desired. ∎

It's particularly interesting to see other ways that this connection given by the distributive property affects what can happen in a ring. Here's another such connection:

THEOREM E.10. *For any $r \in R$ we have that $0_R \cdot r = 0_R$.*

[I] Many texts drop the assumption that a ring contains a multiplicative identity. In doing this one needs to assume E.9 as a part of the definition. As such, to those authors, the even integers, $2\mathbb{Z}$ is a ring, but to us it is not.

Proof. Since 0_R is an additive identity we have that $0_R = 0_R + 0_R$. Therefore, we have

$$r \cdot 0_R = r \cdot (0_R + 0_R) = r \cdot 0_R + r \cdot 0_R$$
$$\Rightarrow 0_R = r \cdot 0_R - r \cdot 0_R = r \cdot 0_R + r \cdot 0_R - r \cdot 0_R = r \cdot 0_R$$

as desired. ∎

Here are a few more examples:

Theorem E.11. *For any ring R:*

(1) $\forall r \in R, (-1_R) \cdot r = r \cdot (-1_R) = -r$

(2) $\forall r, s \in R, (-r) \cdot (-s) = rs \cdot r$

Rant E.12. This is a theorem were students confuse intuition and familiarity. These statements look highly believable and you are certainly used to them in familiar rings like $\mathbb{R}$. However in all of mathematics one of the least intuitive statements is that $(-1) \cdot (-1) = 1$. No, no, no, a statement like "the opposite of an opposite is the original" is nonsense. For instance a dog and a cat are opposites and a cat and a mouse are opposites, but a mouse and dog are certainly not one in the same.[II]

Proof. For starters, it's worth focusing on the fact that -1_R is **by definition** the additive inverse of 1_R. That is, the only thing we know about it is that

$$1_R + (-1_R) = (-1_R) + 1_R = 0_R.$$

(1) We wish to prove that $(-1_R) \cdot r = -r$. This means, by uniqueness of inverses (here negatives, i.e. inverses in the group $(R, +)$) that we need to prove that $(-1_R) \cdot r + r = 0_R$. Indeed

$$\begin{aligned}(-1_R) \cdot r + r &= (-1_R) \cdot r + (1_R) \cdot r \\ &= (-1_R + 1_R) \cdot r && \text{(Right Distrib)} \\ &= 0_R \cdot r && (-1_R \text{ is an additive inverse of } 1_R) \\ &= 0_R.\end{aligned}$$

A similar calculation also proves that $r \cdot (-1_R) + r = 0_R$. Therefore we can conclude the result.

(2) By the first part we know that $-r = (-1_R) \cdot r$ and $-s = (-1_S) \cdot s$. Thus

$$\begin{aligned}(-r) \cdot (-s) &= (-1_R \cdot r) \cdot (-1_R \cdot s) \\ &= -1_R \cdot (r \cdot (-1_R) \cdot s) && (\cdot \text{ is assoc.}) \\ &= -1_R \cdot ((-1_R) \cdot r \cdot s) && \text{(First part)} \\ &= ((-1_R) \cdot (-1_R)) \cdot (r \cdot s) && (\cdot \text{ is assoc.}) \\ &= (-(-1_R)) \cdot (r \cdot s) && \text{(First part)} \\ &= 1_R \cdot (r \cdot s) && \text{(Inverse of inverse is original)} \\ &= r \cdot s\end{aligned}$$

as desired. ∎

[II] Of course in everyday life the word opposite has very little mathematical meaning! Why do people consider cats and dogs to be opposite? They are both so damn cute and fluffy!) Addition of negative numbers is certainly intuitive, if I'm in debt 35 dollars and I get paid 20 dollars I am now in debt 15 dollars, etc. But I would love to hear a reasonable intuition for multiplying such things! If I lose 35 dollars a day and go back in time 3 days would I really have 105 dollars?

EXAMPLE /THEOREM E.13. If R is a ring with $1_R = 0_R$, then R is the zero ring.

PROOF. Suppose that $1_R = 0_R$. Let $r \in R$. We show that $r = 0$. Indeed, $r = r \cdot 1 = r \cdot 0 = 0$. ■

NOTATION E.14. Since the additive part of a ring is just an abelian group when we refer to an adjective which describes a ring we will typically be referring to its multiplicative part. Here are some other definitions and notations (many of which we introduce in the main text but we collect here).

- A *commutative ring* is a ring whose **multiplicative** part is commutative. This means a ring R is *commutative* if for all $r, s \in R$ we have that
$$r \cdot s = s \cdot r$$
- An element $r \in R$ is *invertible* (or a *unit*) if it is invertible in the monoid $(R, \cdot)$. That is if there exists an $s \in R$ so that $r \cdot s = s \cdot r = 1_R$. The set of invertible elements in R is denoted as $\mathcal{U}(R)$.
- An element $r \in R$ is *zero-divisor* if there exists a non-zero $s \in R$ so that $r \cdot s = 0_R$.
- For any ring R (actually the underlying set), the set of all non-zero elements is denoted as $R^\times$.

DEFINITION E.15.

(1) A commutative ring is an *integral domain* if it contains no zero-divisors.

(2) A ring R is a *skew-fields* (or division ring) if $\mathcal{U}(R) = R^\times$.

(3) A commutative skew-field is called a *field*.

In symbols, R is an integral domain if we have that
$$a \cdot b = 0_R \Rightarrow a = 0_R \text{ or } b = 0_R.$$
Recall that every field is an integral domain since an invertible element cannot be a zero-divisor.

EXAMPLE E.16.

(1) $\mathbb{Z}_6$ is a commutative ring which is not an integral domain. This is because $[\,2\,]_6$ is a zero-divisor.

(2) $\mathbb{Z}$ is an integral domain which is not a field. This is because 2 is not invertible.

(3) If D is an integral domain, then so is the polynomial ring $D[x]$. However it is never a field. For instance the polynomial x never has a multiplicative inverse. This is because the product of any polynomial with x has degree at least 1, but the identity 1_R has degree 0.

(4) The quaternions $\mathbb{H}$ form a skew field which is not a field. This is because the multiplication is not commutative, i.e. $\mathbf{ij} \neq \mathbf{ji}$ and every non-zero element is invertible.

(5) For p a positive prime number, $\mathbb{Z}_p$ is a field.

(6) Each of the rings $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are fields.

EXERCISE E.2. Show that

$$\mathbb{Q}[i] \stackrel{\text{def}}{=} \{a + bi : a, b \in \mathbb{Q}\} \quad \text{and} \quad \mathbb{Q}[\sqrt{2}] \stackrel{\text{def}}{=} \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$$

are fields.

SOLUTION

EXERCISE E.3. Prove that

$$\mathcal{U}(\mathbb{Z}[i]) = \{\{n\}_4\}.$$

SOLUTION

Let R be a ring contained inside of the real numbers $\mathbb{R}$. Define $R\mathbb{H}$ as the set of quaternions with coefficients in R. That is

$$R\mathbb{H} \stackrel{\text{def}}{=} \{a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H} : a, b, c, d \in R\}. \tag{1}$$

EXERCISE E.4. Prove for any ring $R \subseteq \mathbb{R}$ that $R\mathbb{H}$ is a ring.

SOLUTION

DEFINITION E.17. Let R be a ring and $r \in R \setminus \mathcal{U}(R)$ (i.e. an non-invertible element of R). The element r is *reducible* if there exists elements s_1 and $s_2 \in R \setminus \mathcal{U}(R)$ with

$$r = s_1 \cdot s_2.$$

Otherwise r is said to be *irreducible*.

Therefore, by definition every element of a ring is either a unit, reducible, or irreducible. For instance, for $R = \mathbb{Z}$ this matches being equal to ± 1, composite, or prime (respectively). If R is a field, like say $\mathbb{Q}$, then every non-zero element is a unit and 0 itself is reducible (since $0 \cdot 0 = 0$). Therefore fields have no reducible elements.

E.2 POLYNOMIALS

You are no doubt familiar with polynomials with coefficients which are real. However we could formally define polynomials where the coefficients are defined in any of the ring. Since such things are more the focus in a second course of abstract algebra, we will not given a precise definition (even in those course such a definition is often skipped!) but we only state that polynomials with coefficients in a ring R defined as formal sums

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots a_n x^n$$

where $a_1, \ldots, a_n \in R$, but $x \notin R$. We call a_i the i^{th} coefficient of $f(x)$, denoted $\text{coef}_i\, f(x) = a_i$. For $i > n$ or $i < 0$ we say that the i^{th} coefficient of $f(x)$ is 0.

RANT E.18. In this way we could formally define a polynomial f as a function $\mathbb{N} \to R$ which eventually equals 0 for large enough numbers. If we denote this function as $f \mapsto f_i$, we can form the what we have been thinking of as a polynomial via

$$f_0 + f_1 x + \cdots f_{n-1} x^{n-1} + f_n x^n$$

where n is the last number so that $f_n \neq 0$. In the other direction you take what we meant as a polynomial and turn it into the function sending a natural number m to to the m^{th} coefficient of the polynomial.

DEFINITION E.19. If $f(x) \neq 0$ define $\deg f(x)$ to be the largest integer i so that $\text{coef}_i f(x) \neq 0$.

CONVENTION E.20. To make some formulas work out better we define the degree of the zero polynomial to be $-\infty$.

Addition and multiplication of polynomials are defined in exactly the same way they are defined for polynomials with real coefficients. Namely given polynomials

$$f(x) = a_0 + \cdots + a_n r^n \text{ and } g(x) = b_0 + \cdots + b_m x^m$$

we define $f(x) + g(x)$ to be the polynomial whose i^{th} coefficient is $a_i + b_i$ and $f(x) \cdot g(x)$ to be the polynomial with coefficients

$$\text{coef}_0(f(x) \cdot g(x)) = a_0 \cdot b_0 \quad \text{coef}_1(f(x) \cdot g(x)) = a_0 \cdot b_1 + a_1 \cdot b_0 \quad \text{coef}_2(f(x) \cdot g(x)) = a_0 \cdot b_2 + a_1 \cdot b_1 + a_2 \cdot b_0$$
$$\cdots \text{coef}_m f(x) \cdot g(x) a_0 b_m + a_1 b_{m-1} + \cdots + a_{m-1} b_1 + a_m b_0 = \sum_{i+j=m} a_i b_j.$$

As such we omit 1's from polynomials. That is we denote the polynomial $1x^3$ simply as x^3. It is easier to use the theorem than the definition to actually perform polynomials multiplications and additions.

EXAMPLE E.21. Over $R = \mathbb{Z}_5$, we have (omitting the class symbols for ease)

$$(1 + 2x^2 + x^5) \cdot (2 + 4x^3 + 2x^5) = (1 \cdot 2 + 1 \cdot 4x^3 + 1 \cdot 2x^5) + (2 \cdot 2x^2 + 2 \cdot 4x^5 + 2 \cdot 2x^7) + (2x^5 + 4x^8 + 2x^{10})$$
$$= 2 + 4x^2 + 4x^3 + 2x^5 + 4x^7 + 4x^8 + 2x^{10}.$$

You can think of this multiplication as taking place in $\mathbb{Z}$ and then taking everything mod n.

THEOREM E.22. *Suppose that R is an integral domain. Then for all $f, g \in R[x]$ we have*

$$\deg(f(x) \cdot g(x)) = \deg f(x) + \deg g(x).$$

PROOF. If

$$f(x) = a_o + a_1 x +$$

■

E.2.1 Division Algorithm for Polynomials

{Fill In }

Theorem E.23. *Suppose that $f(x)$ and $g(x)$ are polynomials with coefficients in a field. Then there exists polynomials $q(x)$ and $r(x)$ so that*

$$f(x) = q(x)g(x) + r(x)$$

where $\deg r(x) < \deg g(x)$ *or* $r(x) = 0$.

Proof. { I guess induction ... } ∎

Theorem E.24. *If x_0 is a root of a polynomial $f(x) \in \mathbb{F}[x]$, then there exists a polynomial $q(x)$ with*

$$\deg g(x) = \deg f(x) - 1 \quad \textit{and} \quad f(x) = g(x) \cdot (x - x_o).$$

Proof. By the Division algorithm for polynomials there exists polynomials $g(x)$ and $r(x)$ so that

$$f(x) = g(x)(x - x_o) + r(x)$$

where $\deg r(x) < \deg (x - x_o) = 1$. Therefore $\deg r(x) = 0$, i.e. $r(x)$ is a constant polynomial, $r(x) = r \in \mathbb{F}$. Then we have

$$0 = f(x_o) = q(x_o)(x_o - x_o) + r = q(x_o) \cdot 0 + r = r.$$

That is $0 = r$. Therefore

$$f(x) = q(x)(x - x_o)$$

as desired. ∎

Corollary E.25. *If $r_1, \ldots, r_n$ are distinct roots of a polynomial $f(x) \in \mathbb{F}[x]$, then there exists some polynomial $q(x)$ of degree* $\deg f(x) - n$

$$f(x) = (x - r_1) \cdots (x - r_n)q(x).$$

and if $f(x)$ has degree n, then

$$f(x) = r\,(x - r_1) \cdots (x - r_n)$$

where $r \in \mathbb{F}$ is the leading coefficient of $f(x)$, i.e. $r = \text{coef}_{\deg f(x)}(f(x))$..

Proof. By the theorem,

$$f(x) = q(x)(x - r_1).$$

We also have

$$0 = f(r_i) = q(r_i)(r_i - r_1).$$

By assumption $r_2 \neq r_1$ and as such $r_2 - r_1 \neq 0$ in this case. Therefore, since $\mathbb{F}$ is a field and as such does not contain any zero divisors $q(r_2)$ must equal 0. Therefore we can apply the theorem to $q(x)$ to factor it as $q(x) = q_2(x)(x - r_2)$ so $f(x) = q_2(x)(x - r_1)(x - r_2)$. Continuing in this way proves the first statement. The second statement follows from comparing degrees and using the fact that

$$(x - r_1) \cdots (x - r_n)$$

has leading term 1. ∎

Theorem E.26. *Let $\mathbb{F}$ be a field. Any finite subgroup of $\mathbb{F}^{\times 1}$ is a cyclic group.*

Proof. Let $n \in \mathbb{Z}$. The equation $x^n = 1$ can have at most n solutions the set

$$\{x : x^n = 1\}$$

has no more than n elements. Therefore for any finite subgroup $C \leqslant \mathbb{F}^\times$ we have

$$\mathbf{O}_n(C) \leq n.$$

Therefore Theorem 13.61 proves that C is cyclic, as desired. ∎

In particular, if $\mathbb{F}$ is finite, then so is $\mathbb{F}^\times$, therefore we have proven

Corollary E.27. *Let $\mathbb{F}$ be a finite field. Then $\mathbb{F}^\times$ is cyclic.* ∎

Special cases of this theorem include:

Corollary E.28.

(1) *Every finite subgroup of $\mathbb{C}^\times$ is cyclic*[II]*;*

(2) *For every prime p, $\mathcal{U}(\mathbb{Z}_p)$ is cyclic.*

Notice this theorem does not hold for the skew field $\mathbb{H}$ (so commutativity plays an essential role); for instance $B\mathcal{D}_9$, $B\left(\blacklozenge^3\right)$, $B\left(\blacktriangleright^3\right)$, and $B\left(\bigstar^3\right)$ are all examples of finite, non-cyclic (not even abelian) subgroups of $\mathbb{H}^\times$.

Exercise E.5. Prove that every finite *abelian* subgroup of $\mathbb{H}^\times$ is cyclic.

SOLUTION

E.2.2 When is a polynomial square-free?

Definition E.29. The *formal derivative* ∇, of a polynomial is defined to be the unique function

$$\nabla\colon \mathbb{F}[x] \to \mathbb{F}[x]$$

with the following 3 properties

PL ∇ is linear. That is for any $c \in \mathbb{F}$ and polynomials f and g we have

$$\nabla(c \cdot f) = c \cdot \nabla(f) \quad \text{and} \quad \nabla(f + g) = \nabla(f) + \nabla(g).$$

PD ∇ is a derivation. That is for any polynomials f and g we have

$$\nabla(f \cdot g) = f \cdot \nabla(g) + \nabla(f) \cdot g.$$

Px ∇ of the polynomial x is 1.

[I] Don't forget, for fields

$$\mathbb{F}^\times = \mathcal{U}(\mathbb{F}).$$

[II] In fact it's equal to $\{\{n\}\}$ for some positive integer n.

Of course we need to prove that such a function exists and is unique..

Exercise E.6. Prove that for all $c \in \mathbb{F}$, $\nabla(c) = 0$.
Hint! First, use the fact that $1 = 1 \cdot 1$ and **PD** to show that $\nabla(1) = 0$. Then, use **PL** to prove the same thing for $c \in \mathbb{F}$.

SOLUTION

Inductively, the definition implies that $\nabla(x^n) = n \cdot x^{n-1}$ (just like the usual derivative). Indeed, this is the case by definition for $n = 1$. Assuming it to be true for n, then

$$\begin{aligned}
\nabla(x^{n+1}) &= \nabla(x \cdot x^n) \\
&= x \cdot \nabla(x^n) + \nabla(x) \cdot x^n && \text{(PD)} \\
&= x \cdot \nabla(x^n) + 1 \cdot x^n && \text{(Px)} \\
&= x \cdot (nx^{n-1}) + x^n && \text{(IH)} \\
&= n \cdot (x \cdot x^{n-1}) + x^n \\
&= nx^n + x^n \\
&= (n+1) \cdot x^n.
\end{aligned}$$

This means that for any polynomial

$$f(x) := a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \cdots + a_1 \cdot x + a_0$$

we have, using this newly derived property and linearity, **PL**,

$$\begin{aligned}
\nabla(f(x)) &= \nabla(a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \cdots + a_1 \cdot x + a_0) \\
&= \nabla(a_n \cdot x^n) + \nabla(a_{n-1} \cdot x^{n-1}) + \cdots + \nabla(a_1 \cdot x) + \cancel{\nabla(a_0)} \\
&= a_n \cdot \nabla(x^n) + a_{n-1} \cdot \nabla(x^{n-1}) + \cdots + a_1 \cdot \nabla(x) \\
&= na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1.
\end{aligned}$$

This is exactly the same formula we'd get if we treated this as a regular everyday derivative. Since we've found a formula for ∇, this proves that ∇ is unique, provided it exists. There's only one step left to do, the boring one:

Exercise E.7. Prove that

$$\widetilde{\nabla}(a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \cdots + a_1 \cdot x + a_0) \stackrel{\text{def}}{=} na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1$$

satisfies all 3 parts of Definition E.29.

SOLUTION

I once had a friend of mine who worked in numerical analysis, and as such didn't give a crap about these formal derivatives, what the point was of defining them at all. His argument was along the lines that "Over some crazy field like $\mathbb{Z}_7$, there is no such thing as a tangent line, or average change, or any sort of limits at all. In fact the formal derivative was not defined in terms of limits as it was in Calculus class, but simply by some properties. How could this thing possibly be useful?"

Here is the first answer to his question:

Proposition E.30. *A polynomial $f(x)$ is square free iff $\nabla(f)$ and f are relatively prime.*

Proof.

$\Rightarrow$ Suppose that $f = g^2 \cdot h$ for some polynomials g and h. Then

$$\begin{aligned}\nabla(f) &= \nabla(g^2 \cdot h)\\ &= g^2 \cdot \nabla(h) + \nabla(g^2) \cdot h\\ &= g^2 \cdot \nabla(h) + (g \cdot \nabla(g) + \nabla(g) \cdot g) \cdot h\\ &= g \cdot (g\nabla(h) + (2h \cdot \nabla(g)))\,.\end{aligned}$$

Therefore $g \mid f$ and $g \mid \nabla(f)$ and so f and $\nabla(f)$ are not relatively prime.

$\Leftarrow$ ∎

E.2.3 Monic Polynomials and Algebraic Integers*

Definition E.31. A polynomial with integer coefficients

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$$

is called *monic* if $a_n = 1$. Roots of monic polynomials are called *algebraic integers*.

It's rather surprisingly difficult to show that the set of all algebraic integers is closed under addition and multiplication in the sense that one needs to introduce a new concept to actually prove it.

Example E.32.

(1) Every integer n is a (the) root of the monic polynomial $x - n$.

(2) The square root of 2, $\sqrt{2}$ is a root of the monic polynomial $x^2 - 2$. Hence $\sqrt{2}$ is an algebraic integer.

(3) The golden ratio, $\phi = \frac{1}{2} + \frac{\sqrt{5}}{2}$ is a root of the monic polynomial $x^2 - x - 1$, so it is an algebraic integer.

(4) More generally, we have that $\frac{1}{2} + \frac{\sqrt{d}}{2}$ is an algebraic integer whenever $d \equiv_4 1$, i.e. $1 - d$ is divisible by 4. Indeed, if we set

$$\alpha = \frac{1}{2} + \frac{\sqrt{d}}{2} \Rightarrow 2\,\alpha = 1 + \sqrt{d} \Rightarrow (2\,\alpha - 1)^2 = d \Rightarrow (4\,\alpha^2 - 4\,\alpha + 1 - d) = 0.$$

Now, the assumption is that $1 - d$ is divisible by 4. Hence $x^2 - x + \frac{1-d}{4}$ is a monic polynomial with α as a root. As such, by definition, α is an algebraic integer.

Using the fact that the product of two algebraic integers is an algebraic integer without proof, we can see that the converse is also true since if $\frac{1}{2} + \frac{\sqrt{d}}{2}$ is an algebraic integer, then it's not hard to see that $\frac{1}{2} - \frac{\sqrt{d}}{2}$ of the same polynomial as the original was. In particular, it is also an algebraic integer. Therefore, their product is also an algebraic integer. But this product is the rational number $\frac{(1-d)}{4}$. But by the proposition just below, this is an algebraic integer only in the case that it's a good old fashioned integer. This happens only when $4 \mid (1 - d)$, i.e. $d \equiv_4 1$ as claimed.

(5) Any n^{th} root of unity is a monic polynomial since it is a root of the monic polynomial $x^n - 1 = 0$.

(6) In a far less trivial example, we prove in Theorem K.7 that for any rational number $\frac{m}{n}$, $2\cos\left(\frac{m}{n}\pi\right)$ is an algebraic integer.

E.3 Homomorphism of Rings

Definition E.33. Let R and S be rings. A function $\varphi\colon R \to S$ is a *homomorphism of rings* if for all r_1 and $r_2 \in R$ we have

RHom 1 (φ **preserves identities**) $\varphi(1_R) = 1_S$,

RHom 2 (φ **preserves addition**) $\varphi(r_1 + r_2) = \varphi(r_1) + \varphi(r_2)$,

RHom 3 (φ **preserves multiplication**) $\varphi(r_1 \cdot r_2) = \varphi(r_1) \cdot \varphi(r_2)$.

The requirement that φ takes the (multiplicative) identity of R to the (multiplicative identity) of S should seem a bit strange since we didn't need to require any such thing for groups. However in many cases we can get away with out it. Here are a few

Example /Theorem E.34.

(1) Suppose that S is an integral domain and φ is not the zero function (i.e. the function which takes everything in R to 0_S). Then we get **RHom 1** for free.

(2) Suppose that φ is surjective. Then we get **RHom 1** for free.

Proof.

(1) First note that $\varphi(1_R) \neq 0_S$. For if it were, then we'd have for all $s \in S$,

$$\varphi(s) = \varphi(s \cdot 1_R) = \varphi(s)\varphi(1_R) = \varphi(s) \cdot 0_S = 0_S$$

which contradicts our assumption that φ is not the zero function.

Next, note that

$$1_S \cdot \varphi(1_R) = \varphi(1_R) = \varphi(1_R \cdot 1_R) = \varphi(1_R) \cdot \varphi(1_R).$$

Subtracting the first term from the last shows that

$$(1_S - \varphi(1_R)) \cdot \varphi(1_R).$$

Therefore since we are assuming that S does not have zero divisors we must have that either $1_S - \varphi(1_R)$ or $\varphi(1_R)$ equals zero. But we already ruled out the latter case, thus the former case must hold. Therefore we have $1_S - \varphi(1_R) = 0_S$ and adding $\varphi(1_R)$ to both sides implies that $1_S = \varphi(1_R)$ as desired.

(2) Now assume that φ is surjective. We show that $\varphi(1_R)$ is 1_S by showing that it satisfies the defining property to the be the identity of $(S, \cdot)$. Since the identity of any algebraic system is unique this will imply what we want. Let $s \in S$. Since φ is assumed to be onto, there exists an $r \in R$ with $\varphi(r) = s$. Then

$$\begin{aligned} \varphi(1_R) \cdot s &= \varphi(1_R) \cdot \varphi(r) \\ &= \varphi(1_R \cdot r) \\ &= \varphi(r) \\ &= s. \end{aligned}$$

Similarly, we also have that $s \cdot \varphi(1_R)$ also equals s. ∎

However it's easy to come up with examples of functions between rings which preserve both operations but do not preserve the identity. For instance the map

$$: \mathbb{Z} \to \mathbb{Z} \times \mathbb{Z}$$
$$: n \mapsto (n, 0)$$

preserves both operations but it does not take the identity of $\mathbb{Z}$ (the number 1) to the identity of $\mathbb{Z} \times \mathbb{Z}$ (the ordered pair $(1, 1)$).

WARNING! E.35. Again, many other text books do not require a ring have a (multiplicative) identity. [I]Therefore they do not require that a homomorphism of rings take the identity to the identity. In this way the map $n \mapsto (n, 0)$ just mentioned would, at least initially, be a perfectly good ring homomorphism.

EXERCISE E.8. Suppose that $\varphi: R \to S$ is homomorphism of rings. Prove

(1) $\varphi(\mathcal{U}(R)) \subseteq \mathcal{U}(S)$

(2) $\varphi|_{\mathcal{U}(R)}: \mathcal{U}(R) \to \mathcal{U}(S)$ is a homomorphism of groups.

(3) If φ is an isomorphism of rings, then $\mathcal{U}(R)$ and $\mathcal{U}(S)$ are isomorphic groups.

(4) By way of counter example, show that the converse of the previous part is false.

SOLUTION

Again, we waste time by making exactly the same definition of the kernel of a homomorphism of rings as we did for the kernel of a homomorphism of groups.

DEFINITION E.36. The kernel of a homomorphism of rings is the set of elements of the domain which get mapped to zero of the codomain.

Therefore, for a ring homomorphism $\varphi: R \to S$, the kernel of φ, denoted $\operatorname{Ker}(\varphi)$ is

$$\operatorname{Ker}(\varphi) \stackrel{\text{def}}{=} \{r \in R : \varphi(r) = 0_S\}.$$

Since a ring homomorphism $R \to S$ is, in particular, a homomorphism between the abelian groups, $(R, +)$ and $(S, +)$, we can point out that the kernel of the homomorphism is the same in either sense. However, I should warn you, there are plenty of group homomorphisms which are not ring homomorphisms! For instance, there are an infinite number of group homomorphism $\mathbb{Z} \to \mathbb{Z}$, but there is only only ring homomorphism, $\mathbb{Z} \to \mathbb{Z}$, namely the identity map.

EXERCISE E.9. Prove this last statement. That is prove that the identity is the only ring homomorphism

$$\mathbb{Z} \to \mathbb{Z}.$$

SOLUTION

[I]This drives me nuts because after one or two chapters they nearly always do require an identity.

Exercise E.10. Prove that the only ring homomorphism $\mathbb{Z}_q \to \mathbb{Z}_q$ is the identity.

SOLUTION

Here is a key property of polynomial rings. In a sense it is the rigorous way of mathematically stating that the variable x can represent any number.

Theorem E.37. *Let $\varphi\colon R \to S$ be a homomorphism of rings. Then for every $s \in S$ there exists a homomorphism $\varphi_s\colon R[x] \to S$ which takes x to s and extends the original, φ.*

Remark E.38. This is without a doubt a theorem which is more obvious than it's proof. The problem is that notation used in the proof (or with any kind of polynomial) rather obstructs the proof.

Proof. The extension is given by "evaluation" in the sense that if

$$f(x) = r_n x^n + r_{n-1}x^{n-1} + \cdots + r_1 x + r_0 \in R[x]$$

and $\varphi(r_i) = s_i$, then define (we bolden s to make it pop)

$$\widetilde{\varphi}(f(x)) = s_n \mathbf{s}^n + s_{n-1}\mathbf{s}^{n-1} + \cdots + s_1 \mathbf{s} + s_0 \in S.$$

For instance for monomials rx^i we are defined $\widetilde{\varphi}(rx^i) = \varphi(r)\mathbf{s}^i$.

It's easy to see that $\widetilde{\varphi}$ satisfies **RHom 1** and **RHom 2**. Moreover when restricted to R (i.e. polynomials without an x term) it coincides with φ. It remains to prove that it also satisfies **RHom 3**. Using the obvious notation, **RHom 3** follows directly from the fact that

$$\left(\sum_{i=0}^{n} s_i \mathbf{s}^i\right)\left(\sum_{i=0}^{m} s'_i \mathbf{s}^i\right) = (s_n \mathbf{s}^n + s_{n-1}\mathbf{s}^{n-1} + \cdots + s_1\mathbf{s} + s_0)\cdot(s'_n \mathbf{s}^n + s'_{n-1}\mathbf{s}^{n-1} + \cdots + s'_1\mathbf{s} + s'_0) \tag{2}$$

$$= \sum_{k=0}^{n+m}(s_0 s'_k + s_1 s'_{k-1} + \cdots s_{k-1}s'_1 + s_k s'_0)\mathbf{s}^k \tag{3}$$

just like for polynomials. In more detail, since φ is a ring homomorphism, we have that

$$\varphi(r_0 r'_k + r_1 r'_{k-1} + \cdots r_{k-1}r'_1 + r_k r'_0) = (s_0 s'_k + s_1 s'_{k-1} + \cdots s_{k-1}s'_1 + s_k s'_0).$$

Indeed, just break up the sum on the left using the fact that φ is a ring homomorphism, replacing r's with s's. Therefore

$$\begin{aligned}\widetilde{\varphi}((r_0 r'_k + r_1 r'_{k-1} + \cdots r_{k-1}r'_1 + r_k r'_0)x^k) &= \varphi(r_0 r'_k + r_1 r'_{k-1} + \cdots r_{k-1}r'_1 + r_k r'_0)\mathbf{s}^k \quad \text{(This is just a monomial)}\\ &= (s_0 s'_k + s_1 s'_{k-1} + \cdots s_{k-1}s'_1 + s_k s'_0)\mathbf{s}^k.\end{aligned}$$

Therefore since $\widetilde{\varphi}$ satisfies **RHom 2**, i.e. preserves addition, we have verified (2). ∎

E.4 Ideals

Definition E.39. Let R be a ring and $I \subseteq R$. I is an ideal (of R) if

(1) I is a subgroup of $(R, +)$.

(2) (**Absorbing Property.**) For any $r \in R$ and $a \in I$ we have that ar and ra are also in I.

If I is an ideal in R we write $I \trianglelefteq R$.

Example E.40.

(1) For any integer n, $n\mathbb{Z} \trianglelefteq \mathbb{Z}$.

(2) The integers are not an ideal in the rationals. For instance the number 1 is an integer and $1/2$ is rational, but $1 \cdot 1/2 = 1/2$ is not an integer.

Proposition E.41. *For an ideal I we have that $I = R$ iff $1_R \in R$.*

We now make the parallel definition of a group generated by a subset.

Definition E.42. Let X be a subset of R. The *ideal generated by X*, denoted $\langle X \rangle$, is an ideal which satisfies

gen 1 $\langle X \rangle \trianglelefteq R$ is an ideal of R.

gen 2 $X \subseteq \langle X \rangle$.

gen 3 If J is an ideal of R which contains X, then $\langle X \rangle \subseteq J$.

The following theorem should be anticipated.

Theorem E.43. *The kernel of any homomorphism of rings is an ideal.*

Proof. Since the kernel of any homomorphism of groups is a subgroup of the domain, we already know that the kernel of a homomorphism of rings (which is a homomorphism of the additive parts with some extra properties) we only need to prove that the kernel is absorbing. To this end let $a \in \mathrm{Ker}\,(\varphi)$ and $r \in R$. We wish to prove that $r \cdot a$ and $a \cdot r$ are in the kernel as well. Indeed

$$\begin{aligned} \varphi(r \cdot a) &= \varphi(r) \cdot \varphi(a) && (\varphi \text{ preserves } \cdot) \\ &= \varphi(r) \cdot 0_S && (a \in \mathrm{Ker}\,(\varphi)) \\ &= 0_S\,. \end{aligned}$$

Thus $r \cdot a$ gets mapped to 0_S, which by its very definition implies that $r \cdot a$ is also in the kernel. The proof that $a \cdot r$ is still in the kernel is the same. ∎

Exercise E.11. Prove that the image of any ring homomorphism

$$\varphi: R \to S$$

we have that $\varphi(R) \leqslant S$.

SOLUTION

Lemma E.44. *The intersection of any collection of ideals is an ideal.*

Proof. Let $\{I_i\}_{i\in\mathcal{A}}$ be a collection of ideals. We already know that the intersection of any collection of subgroups is a subgroup, so we simply need to show that the intersection is also absorbing. To this end let

$$a \in \bigcap_{i\in\mathcal{A}} I_i$$

and $r \in R$. To show that ra is still in this intersection we need to show that $ra \in I_i$ for each $i \in \mathcal{A}$. But since each I_i is assumed to be an ideal, we are given this. ■

Theorem E.45. *For any subset $X \subseteq R$, $\langle X \rangle$ exists and is unique.*

Proof.

Existence Since R is itself an ideal, the collection of all ideals containing X is non-empty. By the lemma this intersection is an ideal. This ideal clearly satisfies both hypothesis needed to qualify as $\langle X \rangle$.

Uniqueness This is the usual argument given several times already. Indeed, if I_1 and I_2 were ideals which satisfied **gen** 1 and **gen** 2, then by the second hypothesis they would both need to be contained inside of each other, and hence be the same. ■

In the case the R is a commutative ring (remember this means the multiplication is commutative!) we can write down what $\langle X \rangle$ is fairly easily.

Definition E.46. Let $r, s \in R$, R a commutative ring. We say that $r \mid s$ if there exists an element $t \in R$ so that $rt = s$.

Proposition E.47. *Let $X \subseteq R$, where R is a commutative ring. Then*

$$\langle X \rangle = \{a_1 r_1 + \cdots a_n r_n : a_i \in R \text{ and } r_i \in X\}.$$

In particular, if X is a single element $X = \{r\}$, then $s \in \langle r \rangle$ iff r divides s.

Proof. Since our goal is to prove that

$$\langle X \rangle = \{r_1 a_1 + \cdots r_n a_n : r_i \in R \text{ and } a_i \in X\}$$

we need to come up with some temporary notation to define the RHS. Let's use (X).

gen 1 Let α and β be arbitrary elements of (X). Then there exists

$$s_1, \ldots, s_n, t_1, \ldots, t_m \in R \quad \text{and} \quad b_1, \ldots, b_n, c_1, \ldots, c_m \in X$$

so that

$$\alpha = \sum_{i=1}^{n} s_i a_i \quad \text{and} \quad \beta = \sum_{i=1}^{m} t_i b_i.$$

Then

$$\alpha + \beta = \sum_{i=1}^{n} s_i a_i + \sum_{i=1}^{m} t_i b_i = \sum_{i=1}^{n+m} r_i a_i$$

where $r_i = s_i$ for $i \leq n$ and $r_i = t_{i-n}$ for $i > n$. Hence this sum is in (X).

Now we show that (X) is absorbing. (Since R is commutative, only need to show left absorbing.) Let $r \in R$ be arbitrary. Then

$$r \cdot \alpha = \sum_{i=1}^{n} r \cdot (s_i a_i) = \sum_{i=1}^{n} (r \cdot s_i) a_i.$$

Since $r \cdot s_i \in R$ we see that $r \cdot \alpha \in (X)$.

gen 2 Now we must show that $X \subseteq (X)$. To this end note that for all $x \in X$, we have that

$$x = 1_R \cdot x.$$

gen 3 Lastly suppose that $J \trianglelefteq R$ and $X \subseteq J$. Let $x_1, \ldots, x_n \in X$ and $r_1, \ldots, r_n \in R$. Since J contains X we have that J also contains all of the $x_i's$. Since J is an ideal it is absorbing. Thus is must contain each of $r_i x_i$'s. Lastly, since an ideal is closed under addition it contains the linear combination

$$r_1 x_1 + \cdots r_n x_n.$$

But an arbitrary element of (X) has this form. Hence $(X) \subseteq J$ as desired.

■

EXAMPLE E.48.

(1) Let R be a commutative ring. Then $\langle x \rangle \trianglelefteq R[x]$ is the set of polynomials without a constant term. Indeed, every such polynomial has the form

$$a_n x^n + \cdots + a_2 x^2 + a_1 x = (a_n x^{n-1} + \cdots a_1) \cdot x.$$

Thus x divides such a polynomial. Conversely, if x divides a polynomial then it clearly cannot have a constant term.

(2) Consider $\langle 2, x \rangle \trianglelefteq \mathbb{Z}[x]$. Then I claim that this is the set of polynomials with even constant term. Indeed, we just explained why every polynomial with zero constant term is in $\langle x \rangle$. Now we can also write every even number as a multiple of 2. Therefore, any polynomail of the form

$$x f(x) + 2n$$

(n an integer, $f(x)$ a polynomial) is in this ideal. Conversely, if $x f(x) + 2g(x)$ is an arbitrary element of this ideal, we write $g(x)$ as $g_1(x)x + c$ where $c = g(0)$ is the constant term of g. Then we'd have

$$x f(x) + 2g(x) = x(x + g_1(x)) + 2c$$

as desired.

DEFINITION E.49. An ideal $I \trianglelefteq R$ is *principle* or *cyclic* if it can be generated by a single element. An integral domain is a *principle ideal domain* (PID) if every ideal is principle.

EXAMPLE E.50.

(1) The integers are a PID. Indeed every subgroup of $\mathbb{Z}$ is cyclic, hence every ideal is as well.

(2) Let $\mathbb{F}$ be a field. Then $\mathbb{F}[x]$ is a PID. Indeed, using exactly the same argument as was used for the integers (i.e. using the division algorithm) one can show that any ideal is generated by a polynomial of least degree in it.

(3) $\mathbb{Z}[x]$ is not a principle ideal domain. Indeed the ideal $\langle 2, x \rangle$ is not a principle ideal, for if it were we'd have a polynomial $f(x) \in \langle 2, x \rangle$ and polynomials $u_1(x)$ and $u_2(x)$ so that $u_1(x) \cdot f(x) = 2$ and $u_2(x) \cdot f(x) = x$. The first equation guarantees that $f(x)$ has degree 0, i.e. is equal to some constant c. Evaluating the second equation at $x = 1$ we see that $u_2(1) \cdot c = 1$, which is an equation in $\mathbb{Z}$. Hence c has a multiplicative inverse in $\mathbb{Z}$ and so $c = \pm 1$, which as a unit, implies that we'd have to have $\langle 2, x \rangle = \mathbb{Z}[x]$. But this is clearly false since, for instance, 1 is not a polynomial with even constant term, and a previous example showed that all elements of $\langle 2, x \rangle$ had this property.

COROLLARY E.51. *Let x be an element of a commutative ring R. Then $\langle x \rangle = R$ iff x is invertible in R.*

PROOF.

$\Leftarrow$ If $\langle x \rangle = R$, then since $\langle x \rangle$ contains everything, it must contain, in particular, 1_R. By the proposition, an element is in $\langle x \rangle$ iff x divides that element. Hence x divides 1_R, in other words x is a unit.

$\Rightarrow$ Suppose that x is a unit, then $\langle x \rangle$ contains a unit, namely x. Hence $\langle x \rangle = R$ as desired. ∎

E.5 QUOTIENT RINGS

One thing that makes ring theory quite appealing is the level of intuition with quotient rings as opposed to quotient groups, particularly in the case of commutative rings.

In this section R is a ring, $I \trianglelefteq R$ and $r, s \in R$.

DEFINITION E.52. As for any group (written additively) the coset $r + I$ is defined to be the set

$$\{r + a : a \in I\}.$$

The set of all cosets of I in R is denoted as R/I, read R mod I, or the quotient of R by I.

DEFINITION E.53. Define $\otimes$ and $\oplus$, : ${}^{R}\!/_{I} \times {}^{R}\!/_{I} \to {}^{R}\!/_{I}$ via

$$(r+I) \otimes (s+I) \stackrel{\text{def}}{=} r \cdot s + I \quad \text{and} \quad (r+I) \oplus (s+I) \stackrel{\text{def}}{=} r + s + I.$$

As we have done many times, we drop the circle around the operations and denote them simply as + and · (or concatenation.)

THEOREM E.54. *The operations $\otimes$ and $\oplus$ are well defined. Moreover*

$$\left({}^{R}\!/_{I},\ \oplus,\ \otimes\right)$$

is a ring under these operations.

PROOF. Thinking of R as a group under addition, the operation $\oplus$ coincides with the operation on the quotient group R/I. Moreover since $(R, +)$ is abelian, $(R/I, \oplus)$ is an abelian group. Thus only need to examine $\otimes$.

(1) (The operation $\otimes$ is well defined) Suppose that $r \equiv_I r'$ and $s \equiv_I s'$. Then $r = r' + a_r$ and $s = s' + a_s$ for some a_r and a_s in I. Therefore

$$\begin{aligned} r \cdot s &= (r' + a_r) \cdot (s' + a_s) \\ &= r' \cdot (s' + a_s) + a_r \cdot (s + a_s) \\ &= r' \cdot s' + r' \cdot a_s + a_r \cdot s + a_r \cdot a_s. \end{aligned}$$

But since I is an ideal, it is absorbing. Thus $r' \cdot a_s \in I$ and $a_r \cdot s \in I$ and $a_r \cdot a_s \in I$. Therefore $r \cdot s - r' \cdot s'$ is the sum of three elements in I. Since I is closed under addition, this difference is also in I. Therefore $r \cdot s \equiv_I r' \cdot s'$, and the operation is well defined.

(2) (**Everything Else**) As with the case of modular arithmetic and quotient groups, it is now elementary to prove that all of the other properties hold. Simply group things in the original ring R and then use the corresponding properties there and break it back down. For instance, to prove associativity of $\otimes$ we have for all $r + I$, $s + I$, and $t + I$ in R/I

$$\begin{aligned} [(r+I) \otimes (s+I)] \otimes (t+I) &= (rs+I) \otimes (t+I) \\ &= (rs)t + I \\ &= r(st) + I \\ &= (r+I) \otimes (st+I) \\ &= (r+I) \otimes [(s+I) \otimes (t+I)] \end{aligned}$$

■

DEFINITION E.55. Let $I \trianglelefteq R$.

(1) I is *maximal* if ${}^{R}\!/_{I}$ is a field.

(2) I is *prime* if ${}^{R}\!/_{I}$ is an integral domain.

THEOREM E.56. *An ideal P is prime iff it satisfies the following:*

$$rs \in P \Leftrightarrow r \in P \text{ or } s \in P. \tag{4}$$

Proof.

⇐ Assume that P is a prime ideal. Suppose that $rs \in P$. Then, we have that

$$rs + P = 0_R + P = 0_{R/P}.$$

But

$$rs + P = (r + P) \cdot (s + P) \Rightarrow (r + P) \cdot (s + P) = 0_{R/P}.$$

However, by definition, since P is prime, R/P is an integral domain. Therefore we must have either $r + P = 0_{R/P}$ or $s + P = 0_{R/P}$, in other words either $r \in P$ or $s \in P$, as desired.

⇒ Suppose that R/P is not an integral domain. Then there would exist r and s so that $(r + P) \cdot (s + P) = 0_R + P$, but neither $s + P$ nor $s + P$ are the zero of the quotient. But just like above, this means that $rs \in P$ but neither r nor s are in P. ■

Theorem E.57. *Let $\mathbb{F}$ be a field and let $f(x) \in \mathbb{F}[x]$. Then* TFAE

(1) *$f(x)$ is an irreducible polynomial,*

(2) *$f(x)$ generates a maximal ideal,*

(3) *$f(x)$ generates a prime ideal.*

E.5.1 Finite Fields

In this exercise, you prove that for any ring R, there is only one ring homomorphism with domain $\mathbb{Z} \to R$

Exercise E.12. Prove for any ring R,

(1) The map

$$\begin{aligned} \chi_R &: \mathbb{Z} \to R \\ &: n \mapsto n \cdot 1_R. \end{aligned}$$

is a ring homomorphism.

(2) If $\mathrm{F} : \mathbb{Z} \to R$ is a ring homomorphism, then $\mathrm{F} = \chi_R$.

Solution

Definition E.58. Let R be a ring. The *characteristic* of R, denoted $\operatorname{Char} R$, is a non-negative generator of the kernel of the ring homomorphism

$$\begin{aligned} \chi_R &: \mathbb{Z} \to R \\ &: n \mapsto n \cdot 1_R. \end{aligned}$$

Proposition E.59. *If* **char** $R = n$, *then for all $r \in R$ we have that $n \cdot r = 0$.*

Proof. We have

$$n \cdot r = n \cdot (1_R \cdot r) = (n \cdot 1_R) \cdot r = 0_R \cdot r = 0_R.$$

■

If we're a bit sloppy, and identify the number n with $n \cdot 1_R$ (which is something that will no doubt be done all of the time in a slightly more advanced text) we have, in a sense, $n = 0_R$ whenever R has characteristic n. Note that the previous proposition and proof are still true for characteristic 0.

Example E.60.

(1) For characteristic not 0, another way of saying this is that $\operatorname{Char} R$ is the least, positive integer n so that $n \cdot 1_R = 0_F$, i.e. it is the additive order of the element $1_R \in R$. Paradoxically, the additive order of 1_R is ∞ iff the characteristic is 0. In any case, the characteristic of R is determined by the additive order of 1_R.

(2) Most rings that you are familiar with before learning any abstract algebra, including $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, have characteristic 0.

(3) The ring $\mathbb{Z}_n$ has characteristic n.

(4) If $S \preccurlyeq R$, then we have, by definition, $1_S = 1_R$. Therefore, since the characteristic of R is determined by the order of 1_R, we have that

$$\operatorname{Char} R = \operatorname{Char} S.$$

Exercise False Theorem E.13. $\operatorname{Char} \mathbb{Z}_9 = 3$

Solution

"**Proof**". Since $\mathbb{Z}_3 \preccurlyeq \mathbb{Z}_9$ we have by Example E.60 **(4)** that

$$\operatorname{Char} \mathbb{Z}_3 = \operatorname{Char} \mathbb{Z}_9.$$

Since $\operatorname{Char} \mathbb{Z}_3 = 3$ we are done. ■

Proposition E.61. *The characteristic of an integral domain is either* 0 *or a prime.*

Proof. (**First we give a fancy, one-line proof.**) Let $\mathbb{D}$ be an integral domain. By definition, the characteristic of $\mathbb{D}$, and by the first isomorphism theorem for rings, we have that $\mathbb{Z}/\operatorname{Char} \mathbb{D} \cdot \mathbb{Z} \cong \mathbb{Z} \cdot 1_{\mathbb{D}}$ is a subring of $\mathbb{D}$. Therefore, as a subring of an integral domain, $\mathbb{Z} \cdot 1_{\mathbb{D}}$ is an integral domain which means that $\operatorname{Char} \mathbb{D}$ is either 0 or a prime, as desired. ■

Proof. (**Another proof of Proposition E.61**) For a more elementary proof, note that if $n > 0$ is not prime, there are integers $1 < q, m < n$ with $n = q \cdot m$ for some $1 < q < n$.

$$n \cdot 1_{\mathbb{D}} = qm \cdot 1_{\mathbb{D}} = (q \cdot 1_{\mathbb{D}})(m \cdot 1_{\mathbb{D}}).$$

Now if n were the characteristic of $\mathbb{D}$ we have that this quantity above is $0_{\mathbb{D}}$. Then as $\mathbb{D}$ is an integral domain at least one of the factors, $q \cdot 1_{\mathbb{D}}$ or $m \cdot 1_{\mathbb{D}}$, must be 0_F. This is impossible since the characteristic is n and $m, q < n$. ■

This gives a way to produce an infinite number of finite fields. For instance, if $f(x) \in \mathbb{Z}_p[x]$ is an irreducible polynomial of degree n then

$$\mathbb{F}_{p^n} \stackrel{\text{def}}{=} \mathbb{Z}_p[x]/\langle f(x) \rangle$$

is a field of order p^n.

Remark E.62. I make no reference to the polynomial in the notation $\mathbb{F}_{p^n}$ since it turns out that up to isomorphism there is a unique field of order p^n for any positive integer n. To see a proof of this (we will not provide one here) check out [DF04] the example starting on the bottom of page 549 titled "Existence and Uniqueness of Finite Fields." Using the result without proof we have that

$$\mathbb{Z}_p[x]/\langle f(x)\rangle \cong \mathbb{Z}_p[x]/\langle g(x)\rangle$$

for any two irreducible polynomials of the same degree.

Proof. Since every non-zero element of $\mathbb{F}$ has additive order p, as an additive group, $\mathbb{F} \cong \mathbb{Z}_p^m$ for some integer m, in particular it means that $|\mathbb{F}| = p^m$. Meanwhile, since $\mathbb{F}^\times$ is the set of units in a field, it must be a cyclic group generated by some $\alpha \in \mathbb{F}^\times$. This means that every non-zero element of $\mathbb{F}$ is a power of α. Let $M(t)$ be the minimal polynomial of α over $\mathbb{Z}_p$, i.e. a polynomial over $\mathbb{Z}_p[t]$ which has α as a root and which has minimal degree with this property. Then we have that

$$\mathbb{Z}_p[t]/\langle M(t)\rangle \cong \mathbb{Z}_p[\alpha] \leqslant \mathbb{F}$$

But since we picked α to be a generator of $\mathbb{F}^\times$, we must have $\mathbb{Z}_p[\alpha] = \mathbb{F}$, therefore $\mathbb{F} \cong \mathbb{Z}_p[t]/M(t)$, and we are done. ∎

By the proof of this we've established

Corollary E.63. *Every finite field has an order equal to the power of some prime.* ∎

As pointed out already,

Proposition E.64. *For any finite field $\mathbb{F}$ we have*

$$\mathbb{F} \cong \mathbb{Z}_p[x]/\langle f(x)\rangle$$

for p the characteristic of $\mathbb{F}$ and $f(x)$ some irreducible polynomial over $\mathbb{F}$. ∎

As pointed out already, one can show moreover, that the field of order p^m is unique up to isomorphism. We won't prove this more general theorem since we won't need it. We'll only use the notation it implies, i.e. we denote any field of order p^m simply as $\mathbb{F}_{p^m}$. We really can only see that this notation is well-defined directly for $m = 2$ (and of course $m = 1$ were $\mathbb{F}_p = \mathbb{Z}_p$) by using the fact that all non-perfect squares differ by a perfect square, i.e. for any two non-perfect squares u_1 and u_2 there exists a $v \in \mathbb{Z}_p$ so that

$$u_1 v^2 = u_2.$$

In characteristic not 2, using the quadratic formula, this guarantees us that every quadratic equation over $\mathbb{Z}_p$

$$ax^2 + bx + c$$

has a root over $\mathbb{Z}_p[u_1]$. In the pesky $p = 2$ situation, this still works since a direct check shows us that the only irreducible polynomial over $\mathbb{Z}_2$ is $x^2 + x + 1$.

means We showed this earlier in this appendix when we showed that

$$\mathbb{Z}_p[\sqrt{u_1}] \cong \mathbb{Z}_p[\sqrt{u_2}].$$

EXAMPLE E.65. The polynomial $x^2 + x + 1$ is irreducible over $\mathbb{Z}_2$. In fact it is the only irreducible polynomial of degree 2 over $\mathbb{Z}_2$. Indeed, there are only 4 degree 2 polynomials over $\mathbb{Z}_2$, 2 of which have constant coefficient 0 (and those sorts of polynomials are divisible by x and so are reducible). The other one, besides $x^2 + x + 1$, is $x^2 + 1$, but this polynomial factors as $x^2 + 1 = (x + 1) \cdot (x + 1)$ over $\mathbb{Z}_2$.Denoting the image of $x \in \mathbb{F}_4$ as ω, the table for $\mathbb{F}_4$ is

$+$	0	1	ω	$\omega+1$
0	0	1	ω	$\omega+1$
1	1	0	$\omega+1$	ω
ω	ω	$1+\omega$	0	1
$\omega+1$	$\omega+1$	ω	1	0

and

$\cdot$	0	1	ω	$\omega+1$
0	0	0	0	0
1	0	1	ω	$\omega+1$
ω	0	ω	$\omega+1$	1
$\omega+1$	0	$\omega+1$	1	ω

EXAMPLE E.66. For primes p other than 2, you can form $\mathbb{F}_{p^2}$ simply by finding and element of $\mathbb{Z}_p$ which is not a perfect square. (Of course we can't do this over $\mathbb{Z}_2$ since the one non-zero element, $[\,1\,]_2$ is a perfect square.) For instance, over $\mathbb{Z}_3$ the element $[\,-1\,]_3$ is not a perfect square (only $[\,1\,]_3$ is a perfect square over $\mathbb{Z}_3$), so the polynomial $x^2 - (-1) = x^2 + 1$ is irreducible over $\mathbb{Z}_3$ and we can take

$$\mathbb{F}_9 = \mathbb{Z}_3[x] \Big/ \left\langle x^2 + 1 \right\rangle.$$

Denoting the image of x in $\mathbb{F}_9$ as i (this is a good notation since we have that $i^2 = -[\,1\,]_2$) we have the following table:

$+$	0	1	-1	i	$i+1$	$i-1$	$-i$	$-i+1$	$-i-1$
0	0	1	-1	i	$i+1$	$i-1$	$-i$	$-i+1$	$-i-1$
1	1	-1	0	$1+i$	$i-1$	i	$-i+1$	$-i-1$	$-i$
-1	-1	0	1	$i-1$	i	$i+1$	$-i-1$	$-i$	
i	i	$i+1$							
$i+1$									
$i-1$									
$-i$									
$-i+1$									
$-i-1$									

and

$+$	0	1	-1	i	$i+1$	$i-1$	$-i$	$-i+1$	$-i-1$
0	0	0	0	0	0	0	0	0	0
1	0	1	-1	i	$i+1$	$i-1$	$-i$	$-i+1$	$-i-1$
-1	0	-1	1	$-i$	$-i-1$	$-i+1$	i	$i-1$	$i+1$
i	0	i	$-i$	-1	$i-1$	$-1-i$	1	$1+i$	$-i-i$
$i+1$	0	$i+1$	$-i-1$	$i-1$	$-i$	1	$-i-1$	-1	i
$i-1$	0	$i-1$	$-i+1$	$-i-1$	1	i	$i-1$	$-i$	1
$-i$	0	$-i$	i	1	$-i+1$	$i+1$	-1	$-i-1$	$i-1$
$-i+1$	0	$-i+1$	$i+1$	$1+i$	-1	$-i$	$-i+1$	i	-1
$-i-1$	0	$-i-1$	$i+1$	$-i+1$	i	-1	$i+1$	1	$-i$

An example of calculation that isn't 100% straight-forward is

$$(i+1)^2 = (1+i) \cdot (1+i) = 1 + i + i + i^2 = 1 + 2 \cdot i - 1 = 2 \cdot i = -i.$$

Note that $\mathbb{F}_9^\times$ is cyclic and generated by $1 + i$ (or by $1 - i$, or by $-1 + i$, or by $-1 - i$). Using this polynomial, the table for $\mathbb{F}_8$ is {do it}

EXAMPLE E.67. The polynomial $f_{31}(x) := x^3 + x + 1$ is irreducible over $\mathbb{Z}_2$. Indeed, as a cubic one can check this simply by noting that it has no roots, i.e. neither $0 = [\,0\,]_2$ nor $1 = [\,1\,]_2$ are solutions of the given polynomial.

Note that the number of reducible degree 3 polynomials is the x^3 coefficient in the taylor expansion of

$$\left(\frac{1}{1-x}\right)^2 \cdot \frac{1}{1-x^2} \tag{5}$$

since is the only way to put together powers of two distinct degree 1 polynomials (this gives rise to the two $\frac{1}{1-x}$ terms) and one degree 2 polynomial, (giving rise to the $\frac{1}{1-x^2}$ term). Doing this directly or cheating and using maple shows us that

$$\left(\frac{1}{1-x}\right)^2 \cdot \frac{1}{1-x^2} = 1 + 2\,x + 4\,x^2 + 6\,x^3 + 9\,x^4 + 12\,x^5 + 16\,x^6 + 20\,x^7 + 25\,x^8 + 30\,x^9 + \cdots .$$

This means there are 6 reducible polynomials of degree 3 over $\mathbb{Z}_2$ and the remaining $2^3 - 6 = 2$ are reducible. Well, we already found one of these two namely $f_{31}(x)$, the other is

$$f_{32}(x) := f_{31}(x+1) = (x+1)^3 + (x+1) + 1 = x^3 + x^2 + x + 1 + x + 1 + 1 = x^3 + x^2 + 1.$$

Example E.68. The polynomial $f_{41}(x) := x^4 + x + 1$ is irreducible over $\mathbb{Z}_2$. One can see that in a number of ways. One way is to note that the only irreducible polynomial of degree 2 over $\mathbb{Z}_2$ is $x^2 + x + 1$ Therefore, if $f_{41}(x)$ were reducible it would be divisible

- x, which it is not since 0 is not a root of $x^4 + x + 1$.
- $x + 1$, which it is not since $-1 = 1$ is not a root of
- $x^2 + x + 1$, which it is not since ω (defined from the example dealing with $\mathbb{F}_4$) is not a root. Indeed,

$$\omega^3 = 1 \Rightarrow \omega^4 = \omega \Rightarrow \omega^4 + \omega + 1 = \omega + \omega + 1 = 0 + 1 = 1 \neq 0.$$

  Alternatively, since $f_{41}(x)$ is not divisible by either x or $x + 1$, if it were reducible it would have to equal

$$(x^2 + x + 1)^2 = x^4 + x^2 + 1,$$

  (a Freshmen's dream was used) which it is not.

(Of course if $f_{41}(x)$ were divisible by a cubic, it would also have to be divisible by a monic, which is something we've already ruled out). Note that the number of reducible degree 4 polynomials is the x^4 coefficient in the taylor expansion of

$$\left(\frac{1}{1-x}\right)^2 \cdot \left(\frac{1}{1-x^2}\right) \cdot \left(\frac{1}{1-x^3}\right)^2$$

where the additional $\frac{1}{1-x^3}^2$ term (compared to the Equation (5)) comes from the 2 irreducible cubic polynomials since is the only way to put together powers of two distinct degree 1 polynomials (this gives rise to the two $\frac{1}{1-x}$ terms) and one degree 2 polynomial, (giving rise to the $\frac{1}{1-x^2}$ term). Expanding the polynomial

$$1 + 2\,x + 4\,x^2 + 8\,x^3 + 13\,x^4 + 20\,x^5 + 31\,x^6 + 44\,x^7 + 61\,x^8 + 84\,x^9 + \cdots .$$

Therefore, there are $2^4 - 13 = 3$ irreducible polynomials of degree 4.

While we can continue on in this fashion indefinitely to compute the number of irreducible polynomials of a given degree over a given $\mathbb{Z}_p$, there is a much more efficient way of doing this using Möbius functions.

Now for any field extension of $\mathbb{Z}_p$, we know that that field extension is of the form

$$\mathbb{Z}_p[x]/\langle f(x) \rangle$$

for some irreducible polynomial $f(x)$. Moreover, if $\deg f = n$, then $\mathbb{Z}_p[x]/\langle f(x) \rangle$ has order p^n. Therefore,

$$\left(\mathbb{Z}_p[x]/\langle f(x) \rangle\right)^\times$$

is a group of order $p^n - 1$. As such, we know that for every non-zero element α of our field extension, we must have that

$$\alpha^{p^n-1} = 1 \Rightarrow \alpha^{p^n} = \alpha.$$

Moreover, we trivially have that $0^{p^n} = 0$. This proves that all p^n elements of $\mathbb{Z}_p[x]/\langle f(x) \rangle$ are roots of the equation $x^{p^n} - x$, which implies that

$$f(x) \Big| \left(x^{p^n} - x\right).$$

Next, note that the formal derivative of $x^{p^n} - x$ is $p^n x^{p^n-1} - 1 = -1$, since the quotient has characteristic p. This is relatively prime with any polynomial, in particular $x^{p^n} - x$ is square free. This proves that

$$x^{p^n} - x = \text{Product of all irreducible poly a divisor of } n.$$

This imp

F Vector Spaces

A vector space over a field $\mathbb{F}$ is an abelian group V and a homomorphism of rings

$$m\colon \mathbb{F} \to \mathrm{End}(V).$$

We write $m(a)(\mathbf{v})$ as $a \cdot \mathbf{v}$.

Given a tuple of vectors $(\mathbf{b}_1, \dots \mathbf{b}_n)$ of a vector space V we can define a linear map $\mathbb{F}^n \to V$ via

$$\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \mapsto c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n.$$

We denote this linear map $(\mathbf{b}_1 | \dots | \mathbf{b}_n)$.

Definition F.1. The span of a set of vectors $\mathcal{B} := \{\mathbf{b}_1, \dots \mathbf{b}_n\} \subseteq V$, denote $\langle \mathcal{B} \rangle$, is defined to be the image of the linear map

$$\mathbf{B} := (\mathbf{b}_1 | \dots | \mathbf{b}_n).$$

$\mathcal{B}$ spans V if $\langle \mathcal{B} \rangle = V$, so that the linear map $\mathbf{B}$ is surjective. The set of vectors are linearly independent if the linear map $\mathbf{B}$ is injective. If $\mathbf{B}$ is both spanning and linearly independent, (so the linear map $\mathbf{B}$ is a bijection) we say that $\mathcal{B}$ is a *basis* for V.

Therefore

$$\langle \mathcal{B} \rangle = \{c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n : c_1, \dots, c_n \in \mathbb{F}\}.$$

Since a linear map is injective iff it has trivial kernel, $\mathcal{B}$ is LI iff we have

$$(\mathbf{b}_1 | \dots | \mathbf{b}_n)\left(\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}\right) = \mathbf{0} \Leftrightarrow \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

In other words $c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n = \mathbf{0}$ iff each $c_i = 0$. So $\mathcal{B}$ is not LI (said *linearly dependent*) if there exists c_i, not all 0, with

$$c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n = \mathbf{0}.$$

If $c_i \neq 0$, then we can solve this equation for $\mathbf{b}_i$, i.e.

$$\mathbf{b}_i = -\frac{1}{c_i}\left(c_1\mathbf{b}_1 + \cdots + c_{i-1}\mathbf{b}_{i-1} + c_{i+1}\mathbf{b}_{i+1} + \cdots + c_n\mathbf{b}_n\right).$$

In other words $\mathbf{b}_i \in \langle\, \mathcal{B} \setminus \{\mathbf{b}_i\}\,\rangle$ and so $\langle\, \mathcal{B} \setminus \{\mathbf{b}_i\}\,\rangle = \langle\, \mathcal{B}\,\rangle$. If $\mathcal{B}$ spans V, then for any $\mathbf{v} \in V$, $\mathcal{B} \cup \{\mathbf{v}\}$ is not linearly independent. If $\mathcal{B}$ is LI (resp. spanning), then so is any subset (resp. superset) of $\mathcal{B}$. If $\mathcal{B}$ is spans, then it must contain a basis.

THEOREM F.2. *Let $\mathcal{B} := \{\mathbf{b}_1, \ldots \mathbf{b}_n\}$ be a basis for V and suppose that $\mathcal{V} := \{\mathbf{v}_1, \ldots \mathbf{v}_m\}$ are linearly independent. Then, $m \leq n$ and there exists $n - m$ elements of $\mathcal{B}$, denoted $\mathcal{B}'$, so that $(\mathcal{B} \setminus \mathcal{B}') \cup \mathcal{V}$ is a basis for V.*

PROOF. Since $\mathcal{B}$ is a basis, it spans. In particular, there exists constants $c_{11}, \ldots, c_{n1}$, so that

$$\mathbf{v}_1 = \sum_{i=1}^{n} c_{i1}\mathbf{b}_i.$$

After permuting $\mathcal{B}$ we can assume that $c_1 \neq 0$. This means that

$$\mathbf{b}_1 = \frac{1}{c}\left(\mathbf{v}_1 - \sum_{i=2}^{n} c_i\mathbf{b}_i\right).$$

In particular, $\mathbf{b}_1 \in \langle\, \mathcal{B}_1\,\rangle$, where $\mathcal{B}_1 := \{\mathbf{v}_1, \mathbf{b}_2, \ldots \mathbf{b}_n\}$. Therefore all of $\mathcal{B}$ are in this span and so $\mathcal{B}_1$ spans V. At the same time this set is LI. {provide} Therefore this set is a basis.

Continuing, we have that $\mathbf{v}_2 \in \langle\, \mathcal{B}_1\,\rangle$ so there are constants $c_{12}, \ldots, c_{n2}$ so that

$$\mathbf{v}_2 = c_{12}\mathbf{v}_1 + \sum_{i=2}^{n} c_{i2}\mathbf{b}_i.$$

If only c_{12} is non-zero we'd have that $\mathbf{v}_2$ is a multiple of $\mathbf{v}_1$. Therefore there exists some $c_{i2} \neq 0$. After permuting $\mathcal{B}^{\geq 2}$ we can assume that $c_{22} \neq 0$. Then follow the exact same process to prove that

$$\mathcal{B}_2 := \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{b}_3, \ldots \mathbf{b}_n\}$$

is a basis.

Continuing in this way, if $m \leq n$, then

$$\mathcal{B}_m := \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m, \mathbf{b}_{m+1}, \ldots, \mathbf{b}_n\}.$$

will form a basis proving the theorem. If $n \leq m$, then we'll end up with a basis

$$\mathcal{B}_n := \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}.$$

If $n < m$, then we'd have that $\mathbf{v}_m$ is not in this set $\mathcal{B}_n$. But since $\mathcal{B}_n$ is a basis it is in $\langle\, \mathcal{B}_n\,\rangle$, which is impossible since $\mathcal{V}$ is assumed to be linearly independent. ■

COROLLARY F.3. *Any two basis of a vector space have the same cardinality.* ■

DEFINITION F.4. This common cardinality is known as the *dimension* of V. It is denoted $\dim V$.

COROLLARY F.5. *If $W \leqslant V$, then* $\dim W \leq \dim V$.

COROLLARY F.6. *Any set of n vectors in a vector space of dimension n which are either LI or spanning form a basis.*

COROLLARY[2] F.7. *A square matrix is injective iff it is surjective.*

PROOF. A way to interpret the previous corollary is by noting that a set of n vectors are LI iff they are spanning. ■

G Using group and ring theory to study matrix groups

In this appendix, I attempt to illustrate to you how one uses both ring and group theory together to prove a number of interesting things about matrix groups. We give to main applications to these results, one in quadratic reciprocity, i.e. the study of when a number is a perfect square in $\mathbb{Z}_p$, and the other application is to the Fibonacci sequence.

G.1 Fields inside $\mathbf{GL}_2\left(\mathbb{Z}_p\right)$

Let p be a prime and $u \in \mathbb{Z}_p$. For shorthand, let's denote the matrix

$$\begin{pmatrix} a & ub \\ b & a \end{pmatrix}$$

as $a + b\,\ulcorner(u)$. We use the typical conventions:

$$\begin{aligned} b\,\ulcorner(u) &\stackrel{\text{def}}{=} 0 + b\,\ulcorner(u) \\ a &\stackrel{\text{def}}{=} a + 0\,\ulcorner(u) \\ a + \ \ulcorner(u) &\stackrel{\text{def}}{=} a + 1\,\ulcorner(u). \end{aligned}$$

These conventions are compatible with matrix arithmetic. For instance the matrix sum of the quantities a and $b\,\ulcorner(u)$ is

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & ub \\ b & 0 \end{pmatrix} = \begin{pmatrix} a & ub \\ b & a \end{pmatrix}$$

which by its very definition $a + b\,\ulcorner(u)$.[I]

Define $\mathbb{Z}_p[\sqrt{u}]$ as the set of such matrices. That is

$$\mathbb{Z}_p[\sqrt{u}] \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} a & ub \\ b & a \end{pmatrix} : a, b \in \mathbb{Z}_p \right\} = \{a + b\,\ulcorner(u) : a, b \in \mathbb{Z}_p\}.$$

Of course (even if $u = 0$) we have a total p^2 elements in $\mathbb{Z}_p[\sqrt{u}]$

Notice that

$$a + b\,\ulcorner(u) + a' + b'\,\ulcorner(u) = (a + a') + (b + b')\,\ulcorner(u)$$

and

$$a + b\,\ulcorner(u) \cdot a' + b'\,\ulcorner(u) = \begin{pmatrix} a & ub \\ b & a \end{pmatrix} \cdot \begin{pmatrix} a' & ub' \\ b' & a' \end{pmatrix} = \begin{pmatrix} aa' + ubb' & u(ab' + a'b) \\ ab' + a'b & aa' + ubb' \end{pmatrix} = (aa' + ubb') + (ab' + a'b)\,\ulcorner(u).$$

[I] These kind of checks may seem odd at first, but note that the plus sign in the original notation $a + b\,\ulcorner(u)$ was meaningless. You need to be careful in math when an author uses notation that seems "too familiar" because it's natural to take certain aspects for granted. For instance we just prove that the + sign in the notation coincides with the addition of matrices.

These equations prove that $\mathbb{Z}_p[\sqrt{u}]$ is a commutative ring whose identity is 1. In fact, for every $u \in \mathbb{Z}_p$ we essentially have just proven that

$$\mathbb{Z}_p[\sqrt{u}] \cong \mathbb{Z}_p[x] \Big/ \langle x^2 - u \rangle.$$

We will not make direct use of this fact, but it may be a useful thing to keep in mind if this jibes with you.

In particular, since

$$(\ulcorner(u))^2 = u \tag{1}$$

we have, in a sense, a square root of u since translating (1) shows us

$$(\ulcorner(u))^2 = u.$$

In fact in later treatments you'll likely see this simply referenced as $\sqrt{u}$ instead of as $\ulcorner(u)$. As this is more than likely the first time you've seen this, we'll stick to our notation because this latter notation may cause confusion. For instance we never have that $\ulcorner(1)$ is not equal to either 1 or -1[II] since $\ulcorner(1)$ is literally the matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

while 1 and -1 are the scalar matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Therefore (provided $p \neq 2$) in the ring $\mathbb{Z}_p[\sqrt{1}]$ we end up having 4 square roots of 1, namely ± 1 and $\pm\ulcorner(1)$.

Using the fact that $(\ulcorner(u))^2 = u$ and the distributive property determines the multiplication in $\mathbb{Z}_p[\sqrt{u}]$ completely:

$$(a + b\,\ulcorner(u)) \cdot (a' + b'\,\ulcorner(u)) = (aa' + bb'u) + (ab' + a'b)\,\ulcorner(u).$$

Lastly, we note that

$$\det(a + b\ulcorner(u)) = \det\begin{pmatrix} a & ub \\ b & a \end{pmatrix} = a^2 - b^2 \cdot u. \tag{2}$$

Since these quantities show up a lot we define

NOTATION G.1.

$$\delta(a + b\,\ulcorner(u)) \stackrel{\text{def}}{=} \det(a + b\,\ulcorner(u)) \text{ and } \overline{a + b\,\ulcorner(u)} \stackrel{\text{def}}{=} a - b\,\ulcorner(u).$$

Notice that as matrix that $\overline{a + b\,\ulcorner(u)}$ is the adjoint of the matrix $a + b\,\ulcorner(u)$. Note also that we have the formula

$$\delta(a + b\,\ulcorner(u)) = \left(a + b\,\ulcorner(u)\right) \cdot \left(\overline{a + b\,\ulcorner(u)}\right).$$

Therefore if $\delta(a + b\,\ulcorner(u)) \neq 0$, then an inverse of $a + b\,\ulcorner(u)$ is

$$\frac{\overline{a + b\,\ulcorner(u)}}{\delta(a + b\,\ulcorner(u))}.$$

On the other hand if $\delta(a + b\,\ulcorner(u)) = 0$, then

$$\left(a + b\,\ulcorner(u)\right) \cdot \left(\overline{a + b\,\ulcorner(u)}\right) = 0$$

and so $a + b\ulcorner(u)$ cannot be invertible (we knew this already since δ coincides with the determinant).

Recall that a non-zero element $\gamma \in \mathbb{Z}_p$ is a *perfect square* if there exists an $x \in \mathbb{Z}_p$ with $x^2 = \gamma$.

[II] Also note that $\ulcorner(0) \neq 0$.

THEOREM G.2. *The element u is is not a perfect square in $\mathbb{Z}_p$ iff $\mathbb{Z}_p[\sqrt{u}]$ is a field.*

PROOF. Since $\mathbb{Z}_p$ is a field, we can test to see whether or not the matrix $a + b\,\ulcorner(u)$ is invertible by checking to see that its determinant is not 0. Clearly, if $b = 0$, then $a + b\,\ulcorner(u)$ is not invertible iff $a = 0$, in which case $a + b\,\ulcorner(u) = 0$. Therefore, we assume that $b \neq 0$.

$\Rightarrow$ Assume that $\mathbb{Z}_p[\sqrt{u}]$ is not a field. Then there exists some $a + b\,\ulcorner(u) \in \mathbb{Z}_p[\sqrt{u}]$ which does not equal 0 and which is not invertible. As pointed out already, we can assume that $b \neq 0$. Then

$$\begin{aligned} &\delta(a + b\,\ulcorner(u)) = 0 \\ &\Rightarrow a^2 - ub^2 = 0 \\ &\Rightarrow a^2 = ub^2 \\ &\Rightarrow \left(\frac{a}{b}\right)^2 = u. \qquad \text{(Used } b \neq 0\text{)} \end{aligned}$$

Hence u is a perfect square.

$\Leftarrow$ Conversely suppose u is a perfect square, i.e. there exists some $x \in \mathbb{Z}_p$ so that $x^2 = u$. Then

$$\delta(x + \ulcorner(u)) = x^2 - u = u - u = 0.^{\text{III}}$$

Therefore, the non-zero (easy to check that x can always be picked so this is the case) element $x + \ulcorner(u)$ is not invertible in $\mathbb{Z}_p[\sqrt{u}]$ and therefore $\mathbb{Z}_p[\sqrt{u}]$ is not a field. ■

REMARK G.3. The intuition behind this is that $\mathbb{Z}_p[\sqrt{u}]$ is in a sense the smallest field which contains both $\mathbb{Z}_p$ and $\sqrt{u}$. If $\mathbb{Z}_p$ happened to already contain a $\sqrt{u}$, then, in a sense, the "overload" causes what results to suck in the sense that things fail to be invertible. As the word "suck" probably gives away, this is just informal reasoning.

This line of reasoning partially explains why the quaternions must be non-abelian. Since the complex numbers already contain the root of any polynomial (i.e. there's not such thing as an irreducible polynomial of degree ≥ 2 in $\mathbb{C}[x]$) (this is the so called *fundamental theorem of algebra* in forming something "bigger" than $\mathbb{C}$ either things must end up becoming non-invertible or things must end up becoming non-abelian.

COROLLARY G.4. *If $u \in \mathbb{Z}_p$ is not a perfect square, then*

$$\left(\mathbb{Z}_p[\sqrt{u}]\right)^\times$$

is a cyclic group of order $p^2 - 1$.

PROOF. By the theorem, $\mathbb{Z}_p[\sqrt{u}]$ is a field. Clearly $\mathbb{Z}_p[\sqrt{u}]$ has p^2 elements. Therefore, by Corollary E.27 $\left(\mathbb{Z}_p[\sqrt{u}]\right)^\times$ is a cyclic group of order $p^2 - 1$. ■

EXAMPLE G.5. (In this example we assume that u is not a perfect square in $\mathbb{Z}_p$) Of course actually finding a generator for $\mathbb{Z}_p[\sqrt{u}]^\times$ is a least as hard as finding a generator for $\mathbb{Z}_p^\times$. Indeed, if $x + y\,\ulcorner(u)$ is a generator for $\left(\mathbb{Z}_p[\sqrt{u}]\right)^\times$, then its $p + 1$ power is a generator for $\mathbb{Z}_p^\times$. Using a Freshmen's Dream, Theorem K.3, that we have

$$(x + y\ulcorner(u))^p = x^p + y^p\ (\ulcorner(u))^p = x + yu^{\frac{p-1}{2}}\ulcorner(u) = x - yu := \overline{x + y\ulcorner(u)}$$

where $u^{\frac{p-1}{2}} = -1$ since u is assumed to not be a perfect square mod p. Therefore

$$(x + y\,\ulcorner(u))^{p+1} = (x + y\,\ulcorner(u))^p \cdot (x + y\,\ulcorner(u)) = (x - y\ulcorner(u)) \cdot (x + y\,\ulcorner(u)) = x^2 - y^2u = \delta(x + y\,\ulcorner(u)).$$

Therefore, we've proven

EXAMPLE /THEOREM G.6. If ω is a generator for $\left(\mathbb{Z}_p[\sqrt{u}]\right)^\times$, then $\delta(\omega)$ is a generator for $\mathbb{Z}_p^\times$. ■

Actually the converse is also true:

EXAMPLE /THEOREM G.7.

For instance, over $\mathbb{Z}_5$ 2 is not a perfect square. Moreover there it turns out that $3 + \ulcorner(u)$ is generator for $\mathbb{Z}_5[\sqrt{2}]^\times$. (Note that $\delta(3 + \ulcorner(u)) = 2$) Indeed, I need to show that this guy has order 24. To do this I just need to show that both $3 + \ulcorner(u)^{12}$ and $3 + \ulcorner(u)^8$ do not equal 1. I illustrate the first one (which is actually easier!) only:

$$\begin{aligned}(3 + \ulcorner(u))^{12} &= \left((3 + \ulcorner(u))^6\right)^2 \\ &= 2^2 \\ &= 4 \\ &\neq 1.\end{aligned}$$

EXERCISE G.1. Prove that $(3 + \ulcorner(u))^8 \neq 1$.

SOLUTION

Having found a generator for the bigger group, namely $\mathbb{Z}_5[\sqrt{2}]^\times$, we can find a generator for the smaller group by computing $(3 + \ulcorner(u))^4$. To do this, we could use a trick:

$$(3 + \ulcorner(u))^4 = (3 + \ulcorner(u))^5 \cdot (3 + \ulcorner(u))^{-1} = (3 - \ulcorner(u)) \frac{(3 - \ulcorner(u))}{2} = (3 - 3\ulcorner(u)) .$$

EXERCISE G.2. Prove that $a + b\ulcorner(u)$ is a generator of $\mathbb{Z}_p[\sqrt{u}]^\times$, then

$$\overline{\frac{a + b\ulcorner(u)}{a + b\ulcorner(u)}}^{\text{IV}}$$

is a generator of $\mathbb{SZ}_p[\sqrt{u}]$.

SOLUTION

G.2 UNDERSTANDING $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ AND $\mathbf{GL}_2\left(\mathbb{Z}_p\right)$

For $u \in \mathbb{Z}_p$, recall the definitions from the last section:

$$\mathbb{Z}_p[\sqrt{u}] := \left\{\begin{pmatrix} a & bub \\ b & a \end{pmatrix} : a, b, u \in \mathbb{Z}_p\right\}.$$

Note to say that elements u and $v \in \mathbb{Z}_p$ differ by a perfect square, i.e.

$$u = a^2 v$$

implies that both u and v are either

[IV] Or expanded out this is the same as

$$\frac{a^2 + b^2 u - 2ab\ulcorner(u)}{a^2 - b^2 u} \quad \text{or} \quad \frac{(a + b\ulcorner(u))^2}{\delta(a + b\ulcorner(u))}.$$

- both perfect squares,
- both not perfect squares, or
- both zero.

THEOREM G.8. *The rings $\mathbb{Z}_p[\sqrt{u}]$ and $\mathbb{Z}_p[\sqrt{v}]$ are conjugate*[I] *iff $\sqrt{u}$ and $\sqrt{v}$ differ by a perfect square.*

PROOF.

$\Leftarrow$ If $u = a^2 v$, then

$$\begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix} \cdot \begin{pmatrix} x & y \cdot u \\ y & x \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}^{-1} = \begin{pmatrix} x & y \cdot \overset{av}{\overset{\shortparallel}{\frac{u}{a}}} \\ ay & x \end{pmatrix} = \begin{pmatrix} x & ay \cdot v \\ ay & x \end{pmatrix},$$

or in terms of our notation

$$\mathrm{D}_{(1,a)} \cdot (x + y \cdot \ulcorner(u)) \cdot \mathrm{D}_{(1,a)}^{-1} = (x + ay \cdot \ulcorner(v)).$$[II]

This proves that $\mathbb{Z}_p[\sqrt{u}]$ and $\mathbb{Z}_p[\sqrt{v}]$ are conjugate.

$\Rightarrow$ To prove the converse, we count the number of non-invertible elements in $\mathbb{Z}_p[\sqrt{u}]$. (Clearly if $\mathbb{Z}_p[\sqrt{u}]$ and $\mathbb{Z}_p[\sqrt{v}]$ are conjugate they have the same number of invertible elements. Why?)

- If u is a perfect square, then, since $\mathbb{Z}_p[\sqrt{u}]$ and $\mathbb{Z}_p[\sqrt{1}]$ and then conjugate we can count the number of non-invertible elements in the latter. But as pointed out the element $(x + y \cdot \ulcorner(1))$ is not invertible iff $x^2 - y^2 = 0$ or $x = \pm y$. Therefore every point in $\mathbb{Z}_p^\times$ contributes two non- invertible elements $(x + x \cdot \ulcorner(1))$ and $(x - x \cdot \ulcorner(1))$ while zero contributes 1, namely $(0 + 0 \cdot \ulcorner(1))$. Therefore there are a total of $2(p-1) + 1 = 2p - 1$ non- invertible elements and

  $$p^2 - (2p - 1) = p^2 - 2p + 1 = (p - 1)^2$$

  invertible elements in $\mathbb{Z}_p[\sqrt{u}]$ for u a perfect square.
- If u is a perfect square, then $\mathbb{Z}_p[\sqrt{u}]$ is a field. Therefore in this case there is 1 non-invertible element and $p^2 - 1 = (p-1)(p+1)$ invertible elements in $\mathbb{Z}_p[\sqrt{u}]$ for u not a perfect square.
- Lastly, if u is 0, then the matrix $(x + y \cdot \ulcorner(0))$ is

  $$\begin{pmatrix} x & 0 \\ y & x \end{pmatrix}$$

  which is not invertible iff $x = 0$. Therefore, in this case there are a total of p non-invertible elements and $p^2 - p = p \cdot (p-1)$ invertible elements in $\mathbb{Z}_p[\sqrt{0}]$.

Since $(p-1)^2$, $(p-1)\cdot(p+1)$, and $p \cdot (p-1)$ can never be the same number these are really three distinct possibilities. ∎

NOTATION G.9. For u a non-square in $\mathbb{Z}_p^\times$, the groups

$$\mathbb{SZ}_p[\sqrt{u}] \overset{\text{def}}{=} \{a + b\sqrt{u} : a^2 - b^2 u = 1\}.$$

[I] This has the obvious meaning. Namely it means there exists an invertible matrix $\mathbf{A}$ with

$$\mathbf{A} \cdot \mathbb{Z}_p[\sqrt{u}] \cdot \mathbf{A}^{-1} = \mathbb{Z}_p[\sqrt{v}].$$

[II] Treating $\ulcorner(u)$ as literally the "square root of u" you'd even expect these two matrices to be equal since when taking square roots of both sides of the equation $u = a^2 v \Rightarrow \sqrt{u} = a\sqrt{v}$ and plugging in would show this. However this is not the case (they are only conjugate) because $\ulcorner(\mathbf{g}\mathrm{tr}\mathbf{A})$ is not literally the square root of $\mathbf{g}$.

Note that in terms of matrices, $\mathbb{SZ}_p[\sqrt{u}]$ is the set of matrices in $\mathbb{Z}_p[\sqrt{u}]$ of determinant 1, i.e.

$$\mathbb{SZ}_p[\sqrt{u}] = \mathbb{Z}_p[\sqrt{u}] \cap \mathbf{SL}_2\left(\mathbb{Z}_p\right).$$

THEOREM G.10. *Suppose $u \in \mathbb{Z}_p$ is not a perfect square. Then,*

$$\mathbb{Z}_p[\sqrt{u}]^\times \quad \textit{and} \quad \mathbb{SZ}_p[\sqrt{u}]$$

are cyclic groups of orders p^2-1 and $p+1$, respectively.

PROOF. Since u is assumed to not be a perfect square, $\mathbb{Z}_p[\sqrt{u}]$ is a field which happens to have order p^2. Therefore, by Corollary E.27, the p^2-1 non-zero elements of $\mathbb{Z}_p[\sqrt{u}]$ forms a cyclic group.

We now tackle $\mathbb{SZ}_p[\sqrt{u}]$. Since $\mathbb{SZ}_p[\sqrt{u}]$ is a subgroup of $\mathbb{Z}_p[\sqrt{u}]^\times$, which we just proved cyclic, $\mathbb{SZ}_p[\sqrt{u}]$ is also cyclic by ??? The last thing we need to prove is

$$\left|\mathbb{SZ}_p[\sqrt{u}]\right| = p+1.$$

To that end consider the function of the homomorphism

$$\begin{aligned} &: \mathbb{Z}_p[\sqrt{u}]^\times \to \mathbb{Z}_p^\times \\ &: a + b\sqrt{u} \mapsto a^2 - b^2 u. \end{aligned}$$

Note that in terms of matrices, this function is a restriction of the determinant function

$$\mathbf{GL}_2\left(\mathbb{Z}_p\right) \to \mathbb{Z}_p^\times$$

and so it is a homomorphism, which we shall denote as $\det|$. Note that by definition, $\left|\mathbb{SZ}_p[\sqrt{u}]\right|$ is the kernel of this homomorphism. Lastly, we need to check that this homomorphism is surjective. Indeed, every perfect square, a^2, in $\mathbb{Z}_p^\times$ is in the image since $\det a + 0\sqrt{u} = a^2$. Since the perfect squares have index 2 in $\mathbb{Z}_p^\times$ we only need to find a single non-perfect square in the image of this map to be able to conclude that $\det|$. From here I'll leave it to you to play an annoying little cat and mouse game similar to the one played in Exercise 13.14 back on page 699. ■

Recall that a scalar matrix is a matrix which is a multiple of the identity matrix, i.e. in the degree 2 case, they are matrices of the form:

$$\begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix}.$$

LEMMA G.11. *Two non-scalar matrices in $\mathbf{GL}_2\left(\mathbb{Z}_p\right)$ are conjugate iff they have the same determinant and trace.*

PROOF.

⇒ Since for matrices $\mathbf{A}$ and $\mathbf{B}$ ($\mathbf{A}$ invertible) we have that

$$\operatorname{tr}(\mathbf{ABA}^{-1}) = \operatorname{tr}(\mathbf{B}) \quad \text{and} \quad \det(\mathbf{ABA}^{-1}) = \det(\mathbf{B})$$

conjugate matrices have the same determinant and trace.

⇐ Suppose $\mathbf{A}$ is a non-scalar matrix. I'm going to show that $\mathbf{A}$ is conjugate to the matrix

$$\mathbf{A}' := \begin{pmatrix} 0 & -\det\mathbf{A} \\ 1 & \operatorname{tr}\mathbf{A} \end{pmatrix}.$$

Indeed, the assumption that $\mathbf{A}$ is non-scalar means there exists a $\mathbf{v}$ so that $\mathbf{A}(\mathbf{v})$ is not a multiple of $\mathbf{v}$. This means that the linear map $\mathbf{T}$ defined by taking $\mathbf{e}_1$ to $\mathbf{v}$ and $\mathbf{e}_2$ to $\mathbf{A}(\mathbf{v})$ is invertible. In order to prove that $\mathbf{A} = \mathbf{T} \cdot \mathbf{A}'\mathbf{T}^{-1}$ I can show that

$$\mathbf{A} \cdot \mathbf{T} = \mathbf{T} \cdot \mathbf{A}'.$$

Indeed, this is a direct calculation:

$$(\mathbf{A}\cdot\mathbf{T})(\mathbf{e}_1) = \mathbf{A}(\mathbf{v}) \quad \text{and} \quad (\mathbf{A}\cdot\mathbf{T})(\mathbf{e}_2) = \mathbf{A}(\mathbf{A}(\mathbf{v})) = \mathbf{A}^2(\mathbf{v}) = \mathrm{tr}\mathbf{A}\mathbf{A}(\mathbf{v}) - \det\mathbf{A}\mathbf{v}.$$

where we have made use of Cayley-Hamilton. On the other hand, seeing what the RHS does to $\mathbf{e}_1$ and $\mathbf{e}_2$:

$$(\mathbf{T}\cdot\mathbf{A}')(\mathbf{e}_1) = \mathbf{T}(\mathbf{e}_2) = \mathbf{A}(\mathbf{v}) \quad \text{and} \quad (\mathbf{T}\cdot\mathbf{A}')(\mathbf{e}_2) = \mathbf{T}(\mathbf{e}_2) = \mathbf{T}(\mathrm{tr}\mathbf{A}\mathbf{e}_2 - \det\mathbf{A}\mathbf{e}_1) = \mathrm{tr}\mathbf{A}\mathbf{A}(\mathbf{v}) - \det\mathbf{A}\mathbf{v}.$$

Since both sides do the same thing to the standard basis vectors, both sides are equal. Therefore $\mathbf{A}$ is conjugate to $\mathbf{A}'$. Since $\mathbf{A}'$ depends only on the trace and determinant of the matrix $\mathbf{A}$, any two matrices of the same trace and determinant are conjugate to the same matrix, hence are conjugate to each other. ■

WARNING! G.12. We have proven that any two such matrices are conjugate over $\mathbf{GL}_2\left(\mathbb{Z}_p\right)$. However this is false if we say that the marices are conjugate over $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$. For instance, over $\mathbb{Z}_3$ it's not difficult to show that the matrices

$$\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$$

are **not** conjugate to each other. (The theorem only guarantees that they are conjugate to each other over $\mathbf{GL}_2(\mathbb{Z}_3)$). However, in a sense, this is the only thing that can go wrong over $\mathbf{SL}_2$ ().

NOTATION G.13. Let $\mathbf{A} \in \mathbf{M}_2(\mathbb{F})$. Define $\gamma(\mathbf{A})$ to be

$$\mathrm{tr}(\mathbf{A})^2 - 4\det\mathbf{A}.$$

For those who've taken Linear Algebra, you should note that $\gamma(\mathbf{A})$ is the discriminant (i.e. the stuff under the square root in the quadratic formula) of the characteristic polynomial of the 2 by 2 matrix $\mathbf{A}$:

$$x^2 - \mathrm{tr}(\mathbf{A})\,x + \det(\mathbf{A}).$$

EXERCISE G.3. Prove that for all $\mathbf{A} \in \mathbf{GL}_2(\mathbb{Z}_2)$, $\gamma(\mathbf{A}^2) = \mathrm{tr}(\mathbf{A})^2 \cdot \gamma(\mathbf{A})$. Conclude that $\gamma(\mathbf{A}^2)$ is a perfect square iff $\gamma(\mathbf{A})$ is.

SOLUTION

Suppose that $\mathbb{F}$ is a field and that $\mathbf{A} \in \mathbf{SL}_2(\mathbb{F})$. There are three cases to consider:

Case 1 If $\gamma(\mathbf{A})$ is a perfect square in $\mathbb{F}$. (Often called the *hyperbolic* case.)

Case -1 If $\gamma(\mathbf{A})$ is not a perfect square in $\mathbb{F}$. (Often called the *elliptic* or *spherical* case.)

Case 0 If $\gamma(\mathbf{A}) = 0$. (Often called the *parabolic* or *flat* case.)

RANT G.14. In the case that $\mathbb{F} = \mathbb{R}$, the real numbers, for matrices of determinant 1, the only way a matrix fits into **Case 1** is if $|\mathrm{tr}(\mathbf{A})| > 2$ in which case such a matrix cannot possibly have finite order. The only way a matrix fits into **Case -1** is if $|\mathrm{tr}(\mathbf{A})| < 2$. The distinction between the three types of matrices in this case is often the subject of a basic course on Dynamical systems or in a course on hyperbolic geometry. Indeed, such matrices essentially make up the isometry group of the hyperbolic plane (actually the isometry group of the hyperbolic plane is $\mathbf{PSL}_2(\mathbb{R})$) and matrices that fall into **Case 1** fix two points at infinity, matrices which fall into **Case -1** fix one point in the actual plane, and matrices which fall into **Case 0** fix exactly one point at infinity. { insert reference to Boris' book}

Notice that in **Case 0**, if we moreover assume that $\mathbf{A}$ has determinant 1, then we must have that $\operatorname{tr}(\mathbf{A}) = \pm 2$. Of course any scalar matrix $\lambda \mathbf{I}_2$ falls into **Case 0**, but many non-scalar matrices like

$$\mathfrak{Q} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

do as well.

For the rest of the section we consider only the case that $\mathbb{F} = \mathbb{Z}_p$ where p is an odd prime, i.e. a prime other than 2.

THEOREM G.15. *Let $\mathbf{A} \in \mathbf{GL}_2(\mathbb{Z}_p)$. Denote t as $\operatorname{tr}(\mathbf{A})$ and γ as $\gamma(\mathbf{A})$. Then for*

(1) $\mathbf{A}$ *in* **Case 1***:*

- $\mathbf{A}$ *is conjugate in* $\mathbf{GL}_2(\mathbb{Z}_p)$ *to*

$$\frac{1}{2}\begin{pmatrix} t+\sqrt{\gamma} & 0 \\ 0 & t-\sqrt{\gamma} \end{pmatrix}$$ [III]

- *The order of* $\mathbf{A}$ *is* $\left|\frac{1}{2}\left(t+\sqrt{\gamma}\right)\right|_{\cdot}$, *and divides* $p-1$*;*
- *The centralizer of* $\mathbf{A}$ *in* $\mathbf{SL}_2(\mathbb{Z}_p)$ *is isomorphic to* $\mathbb{Z}_{p-1}$ (*hence cyclic*), *and the centralizer in* $\mathbf{GL}_2(\mathbb{Z}_p)$ *is isomorphic to* $\mathbb{Z}_{p-1} \times \mathbb{Z}_{p-1}$ (*hence* **not** *cyclic*).

(2) $\mathbf{A}$ *in* **Case -1***:*

- $\mathbf{A}$ *is conjugate in* $\mathbf{GL}_2(\mathbb{Z}_p)$ *to*

$$\frac{1}{2}(t + v^{\Gamma}(\gamma)) = \frac{1}{2}\begin{pmatrix} t & \gamma \\ 1 & t \end{pmatrix}.$$

  (*Which is an element of* $\mathbb{Z}_p[\sqrt{\gamma}]$.)
- *If* $\det \mathbf{A} = 1$, *this order divides* $p+1$*;*
- *The centralizer of* $\mathbf{A}$ *in* $\mathbf{SL}_2(\mathbb{Z}_p)$ *is isomorphic to* $\mathbb{Z}_{p+1}$ (*hence cyclic*), *and in* $\mathbf{GL}_2(\mathbb{Z}_p)$ *it is isomorphic to* $\mathbb{Z}_{p^2-1}$ (*hence cyclic*).

(3) $\mathbf{A}$ *in* **Case 0** *and* $\mathbf{A} \neq \pm \mathbf{I}_2$*:*

- $\mathbf{A}$ *is conjugate in* $\mathbf{GL}_2(\mathbb{Z}_p)$ *to*

$$\frac{1}{2}\begin{pmatrix} t & 0 \\ 1 & t \end{pmatrix}.$$ [IV]

- $|\mathbf{A}| = p \cdot \left|\frac{1}{2}t\right|_{\cdot}$.
- *The centralizer of* $\mathbf{A}$ *in* $\mathbf{SL}_2(\mathbb{Z}_p)$ *is isomorphic to* $\mathbb{Z}_{2p}$ (*hence cyclic*), *and in* $\mathbf{GL}_2(\mathbb{Z}_p)$ *it is isomorphic to* $\mathbb{Z}_{p\cdot(p-1)}$ (*hence cyclic*).

[III] Having fixed a square root, $\sqrt{\gamma}$ of what we're assuming is a perfect square, γ.
[IV] Note that **Case 3** fits into case **Case 2** with $\gamma = 0$. However, recall that the $\mathbb{Z}_p[\sqrt{a}]$ are only fields for $a \in \mathbb{Z}_p$ not a perfect square (and not zero).

Proof. Checking every case, we can see that the trace in every case is the trace of the original, t, and the determinant of the proposed matrices in every case is also the same as the original, i.e. 1. The trace check is obvious since in every case we're checking against a matrix whose diagonals are either both t (in the second two cases) or one diagonal entry is t + something and the other is t − something (in the first case). In each case the determinant checks come down to the relation

$$\frac{1}{2}\left(t^2 - \gamma\right) = 1.$$

which holds by definition of γ.

The orders of the elements comes down to checking case by case the orders of the representative elements. Note in the second case, we are conjugate to

$$\frac{1}{2}\left(t + \ulcorner(\gamma)\right)$$

in $\mathbb{SZ}_p[\sqrt{\gamma}]$, which is a cyclic group of order $p+1$.

Lastly, the statement about centralizers again comes down to checking the orders of the centralizer of the representative elements.

Case 1 Use the notation:

$$\mathbf{A}' := \frac{1}{2}\begin{pmatrix} t+\sqrt{\gamma} & 0 \\ 0 & t-\sqrt{\gamma}. \end{pmatrix}$$

In the notation of Example/Theorem 6.40 back on page 336

$$\mathbf{A}' = \mathbf{D}_{\frac{1}{2}(t+\sqrt{\gamma})}.$$

Note that that Example/Theorem is applicable since

$$t+\sqrt{\gamma} \neq t-\sqrt{\gamma}$$

since $2\sqrt{\gamma} \neq 0$ because p is assumed odd (i.e. $p \neq 2$) so $2 \neq 0$ and $\sqrt{\gamma} \neq 0$ by assumption. Therefore, using that example/theorem, we can conclude that the centralizer of $\mathbf{A}'$ is the set of matrices of the form:

$$\mathbf{D}_\lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}.$$

Therefore the centralizer of $\mathbf{A}'$ is isomorphic to $\mathbb{Z}_p^\times$, a cyclic group of order $p-1$. Since $\mathbf{A}$ and $\mathbf{A}'$ are conjugate this also implies that our original $\mathbf{A}$ has the same property: i.e. a cyclic centralizer of order $p-1$.

Case -1 Of course any two matrices in $\mathbb{Z}_p[\sqrt{\gamma}]$ commute with each other (in fact this set formed a field). For

$$a + b\ulcorner(\gamma) \in \mathbb{Z}_p[\sqrt{\gamma}] \text{ with } b \neq 0$$

(the given matrix is of this form with $a = \frac{t}{2}$ and $b = \frac{1}{2}$) we show that this is the entire centralizer. For any $\mathbf{A} \in \mathbf{GL}_2\left(\mathbb{Z}_p\right)$ we have

$$\mathbf{A}\left(a + b\ulcorner(\gamma)\right)\mathbf{A}^{-1} = a + b\mathbf{A}\ulcorner(\gamma)\mathbf{A}^{-1}.$$

Therefore, $\mathbf{A}$ commutes with $a + b\ulcorner(\gamma)$ iff

$$a\mathbf{I}_2 + b\mathbf{A}\ulcorner(\gamma)\mathbf{A}^{-1} = a\mathbf{I}_2 + b\ulcorner(\gamma) \Leftrightarrow \mathbf{A}\ulcorner(\gamma)\mathbf{A}^{-1} = \ulcorner(\gamma)$$

i.e. iff $\mathbf{A}$ commutes with $\ulcorner(\gamma)$. A direct check shows that for an arbitrary matrix

$$\mathbf{A} := \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

we have

$$\sqrt{\gamma}\cdot\mathbf{A}\cdot\sqrt{\gamma}^{-1} = \begin{pmatrix} 0 & \gamma \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix}\begin{pmatrix} 0 & \gamma \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} w & z\gamma \\ y\gamma^{-1} & x \end{pmatrix}.$$

As such, $\mathbf{A}$ commutes with $\ulcorner(\gamma)$ iff $x = w$ and $zu = y$. Therefore,

$$\mathbf{A} = \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} x & z\gamma \\ z & x \end{pmatrix}.$$

This happens to equal the matrix $x + z\ulcorner(\gamma)$. Hence $\mathbf{A} \in \mathbb{Z}_p[\sqrt{\gamma}]$ proving that $\mathbf{A}$ commutes with $a + b\ulcorner(\gamma)$ iff $\mathbf{A} \in \mathbb{Z}_p[\sqrt{\gamma}]$. Taking appropriate intersections establishes that (technically, this is the first time we're actually using the assumption that γ is not a perfect square in $\mathbb{Z}_p$.)

$$C_{\mathbf{GL}_2(\mathbb{Z}_p)}(a + b\sqrt{\gamma}) = \left(\mathbb{Z}_p[\sqrt{\gamma}]\right)^\times \quad \text{and} \quad C_{\mathbf{SL}_2(\mathbb{Z}_p)}(a + b\sqrt{\gamma}) = \mathbb{SZ}_p[\sqrt{\gamma}].$$

(These are cyclic groups of orders $p^2 - 1$ and $p + 1$ respectively.)

Case 0 In the third case you have that

$$C_{\mathbf{GL}_2(\mathbb{Z}_p)}(\mathbf{A}') = \left\{\begin{pmatrix} a & b \\ 0 & a \end{pmatrix} : a \in \mathbb{Z}_p^\times,\ b \in \mathbb{Z}_p\right\} \quad \text{and} \quad C_{\mathbf{SL}_2(\mathbb{Z}_p)}(\mathbf{A}') = \left\{\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} : b \in \mathbb{Z}_p\right\} \cup \left\{\begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix} : b \in \mathbb{Z}_p\right\}.$$

The latter is a cyclic group generated by

$$\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}.$$

Indeed, this calculation was done in Chapter 6, in Example 6.38 back on page 330.

Therefore, the result has been proven in all cases. ■

REMARK G.16. We could have given a more symmetric but slightly less useful theory to all of this. Namely, we have that every matrix $\mathbf{A} \in \mathbf{M}_2\left(\mathbb{Z}_p\right)$ is conjugate to a matrix in $\mathbb{Z}_p[\sqrt{\gamma(\mathbf{A})}]$. Since Theorem G.8 implies that $\mathbb{Z}_p[\sqrt{u}]$ and $\mathbb{Z}_p[\sqrt{\gamma(\mathbf{A})}]$ are conjugate iff

- both perfect squares,
- both not perfect squares, or
- both zero.

this would means if we fix our favorite non-perfect square in $\mathbb{Z}_p$ (if -1 is available, I like to use this, if not it doesn't much matter which you pick. Of course for a perfect square we always can, and will, choose the best perfect square of them all, namely 1) every matrix, depending on whether $\gamma(\mathbf{A})$ is conjugate to a matrix in $\mathbb{Z}_p[\sqrt{u}]$, $\mathbb{Z}_p[1]$, or $\mathbb{Z}_p[0]$ (the only thing these three rings have in common are just the scalar matrices). However, in our treatment, since in itself the ring $\mathbb{Z}_p[1]$ doesn't appear to be overly friendly, we replaced it with diagonal matrices which are much easier to deal with.

EXAMPLE G.17. To be a jerk, in Exercise 6.25 back on page 335 I gave you an exercise which asked you to compute the centralizer in $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ of the cat matrix

$$\mathbf{C} := \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

where p is

(1) $\mathbf{SL}_2(\mathbb{Z}_3)$ **(2)** $\mathbf{SL}_2(\mathbb{Z}_{11})$.

Doing this directly probably turned into a bit of a slog. But armed with our awesome new theorem this is trivial (note that $\gamma(\mathbf{C}) = 3^2 - 4 = 5$).

(1) Since 5 is not a perfect square mod 3, this means that over $\mathbb{Z}_3$ you would have ended up proving that the order of this centralizer, in this case, is $3 + 1 = 4$.

(2) Since 5 is a perfect square mod 11 (this can be seen by using the lemma below or directly since $4^2 \equiv_{11} 5$) you would have ended up proving that the order of this centralizer, in this case, is $11 - 1 = 10$.

COROLLARY G.18.

(1) *Every abelian subgroup of* $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ *is cyclic.*

(2) *The centralizer of every non-scalar matrix in* $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ *is cyclic and is abelian in* $\mathbf{GL}_2\left(\mathbb{Z}_p\right)$.

(3) *There exists an element of order* n *in* $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ *iff* $n = p, 2p$, *or if* n *divides* $p - 1$ *or* $p + 1$.

LEMMA G.19. *There exists a matrix* $\mathbf{A} \in \mathbf{SL}_2\left(\mathbb{Z}_p\right)$ *of order* 12 *iff* 3 *is a perfect square in* $\mathbb{Z}_p$.

PROOF. The matrix $\mathbf{A}$ has order 12 iff $\mathbf{A}^2$ has order 6. But a as matrix in $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ this means that $\operatorname{tr}\left(\mathbf{A}^2\right) = 1$. But using Cayley-Hamilton we know that

$$\operatorname{tr}\left(\mathbf{A}^2\right) = \operatorname{tr}(\mathbf{A})^2 - 2 \Rightarrow \operatorname{tr}(\mathbf{A})^2 = 3.$$

Therefore, a matrix has order 12 iff its trace is a square root of 3. Since this trace is in $\mathbb{Z}_p$ we have the result. ■

PROPOSITION G.20. *The element* $[3]_p \in \mathbb{Z}_p$ *is a perfect square iff* $p \equiv_{12} 1$ *or* $p \equiv_{12} -1$.

PROOF. By the corollary, Corollary G.18, $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ contains an element of order 12 iff $12 \mid (p-1)$ or $12 \mid (p+1)$, i.e. iff $p \equiv_{12} 1$ or $p \equiv_{12} -1$, as desired. ■

This was essentially just a warm up. Due to its important in the Fibonacci sequence, the number we care a lot more about is 5. Note that $\mathbb{Z}_p$ contains a $\sqrt{5}$ iff $\mathbb{Z}_p$ contains a golden ratio:

$$\frac{1}{2} + \frac{\sqrt{5}}{2},$$ [V]

i.e. an $\phi \in \mathbb{Z}_p$ which satisfies $\phi^2 = \phi + 1$.

PROPOSITION G.21. $\mathbb{Z}_p$ *contains a* $\sqrt{5}$ *iff* $p \equiv_5 1$ *or* $p \equiv_5 -1$. [VI]

PROOF. The question does not make sense for $p = 5$, so we assume that $p \neq 5$. We need to prove that $\mathbb{Z}_p$ contains a $\sqrt{5}$ iff $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ contains an element of order 5. Actually, since 5 is odd, this is equivalent to finding a matrix $\mathbf{A}$, other than $-\mathbf{I}_2$, with $\mathbf{A}^5 = -\mathbf{I}_2$ (simply take the negative of a matrix of order 5). As pointed out already, $\mathbb{Z}_p$ contains a $\sqrt{5}$ iff it contains a golden ratio ϕ. Consider the equation $\mathbf{A}^5 + \mathbf{I}_2 = 0$ which factors as

$$(\mathbf{A} + \mathbf{I}_2) \cdot (\mathbf{A}^4 - \mathbf{A}^3 + \mathbf{A}^2 - \mathbf{A} + \mathbf{I}_2) = 0. \tag{3}$$

[V] Again, we need to make use of our assumption that $p \neq 2$.

[VI] Since the only perfect squares in $\mathbb{Z}_5$ are 1 and -1, a fancier way of stating the proposition is $\mathbb{Z}_p$ *contains a square root of* 5 *iff* $\mathbb{Z}_5$ *contains a square root of* p.

Now the matrix $(\mathbf{A}+\mathbf{I}_2)$ has determinant $2+\text{tr}\mathbf{A}$, so it is invertible since a trace 2 matrix cannot possibly have order 5 since trace 2 matrices always have order p and we have $p \neq 5$. Therefore we can cancel out the $(\mathbf{A}+\mathbf{I}_2)$ term in 3 to simplify us down to

$$(\mathbf{A}^4-\mathbf{A}^3+\mathbf{A}^2-\mathbf{A}+\mathbf{I}_2)=0 \Rightarrow \mathbf{A}^2-\mathbf{A}+\mathbf{I}_2-\mathbf{A}^{-1}+\mathbf{A}^{-2}=0 \tag{4}$$

which we obtain by multiplying both sides by $\mathbf{A}^{-1}$. But in general, for a matrix $\mathbf{B}$ of determinant 1, we always have

$$\mathbf{B}+\mathbf{B}^{-1}=\text{tr}(\mathbf{B})\,\mathbf{I}_2$$

which allows Equation (6) to simplify further to

$$\text{tr}\left(\mathbf{A}^2\right)\cdot\mathbf{I}_2-\text{tr}(\mathbf{A})\cdot\mathbf{I}_2+\mathbf{I}_2=0. \tag{5}$$

Using Cayley-Hamilton to plug in $\text{tr}\left(\mathbf{A}^2\right)=\text{tr}(\mathbf{A})^2-2$ into this equation gives us

$$\text{tr}(\mathbf{A})^2\cdot\mathbf{I}_2-\text{tr}(\mathbf{A})\cdot\mathbf{I}_2-\mathbf{I}_2=0 \Rightarrow \text{tr}(\mathbf{A})^2\cdot\mathbf{I}_2=\text{tr}(\mathbf{A})\cdot\mathbf{I}_2+\mathbf{I}_2=0 \Rightarrow \text{tr}(\mathbf{A})^2=\text{tr}(\mathbf{A})+1. \tag{6}$$

In other words, $\text{tr}(\mathbf{A})$ must be golden ratio. ■

The definition of the Fibonacci sequence, F_n, while simple, always seems rather random to me. That's why I'm always so impressed by rather shocking results like the following one which even further refine the rather impressive results from Exercise 11.17 back on page 622.

THEOREM G.22. *Let p be a prime number other than* 5. *Then $p \mid F_{p-1}$ iff $p \equiv_5 \pm 1$ and $p \mid F_{p+1}$ iff $p \equiv_5 \pm 2$.*

PROOF. Since $F_3=2$, the theorem is true for the special case that $p=2$. Therefore we can assume, in addition that $p \neq 5$, that $p \neq 2$ as well.

Let Φ be the golden matrix. In the exercise referenced before the theorem, Exercise 11.17, you showed that $p \mid F_n$ iff $[\,\Phi\,]_p^n$ is a diagonal matrix (coefficients taken in $\mathbb{Z}_p$). Even though the determinant of the golden matrix is not 1, it's -1, the determinant of the cat matrix $\mathbf{C}:=\Phi^2$ is 1. Note that $\gamma(\mathbf{C})=5$, and the assumption that p does not equal 5 means that $\mathbf{C}$ (and Φ) does not fall into . Therefore, by the results of this section we know that $\mathbf{C}$ has order which divides either $p-1$ or $p+1$, respectively depending on whether or not $\gamma(\mathbf{C})=5$ is a perfect square in $\mathbb{Z}_p$ or not. The only cause for concern is that we're dealing with $\mathbf{C}$ and not the original golden matrix Φ. But since p is odd, both $p+1$ and $p-1$ are even and both Φ^{p-1} and Φ^{p+1} have determinant 1. Therefore, if say

$$\mathbf{C}^{p-1}=\mathbf{I}_2 \Rightarrow \left(\Phi^{p-1}\right)^2=\mathbf{I}_2 \Rightarrow \Phi^{p-1}=\pm\mathbf{I}_2$$

where we've used the fact that the only determinant 1 matrices, which (again) Φ^{p-1} is an example of, which square to the identity are $\pm\mathbf{I}_2$. In particular, Φ^{p-1} is a diagonal matrix and in this case $p \mid F_{p-1}$. The case that $(\mathbf{C})^{p+1}$ similarly gives us the conclusion that $p \mid F_{p+1}$. In conclusion, if 5 is a perfect square in $\mathbb{Z}_p$, then $p \mid F_{p-1}$ and if it is not, then $p \mid F_{p+1}$. The proof is finished after we use Proposition G.21 to classify when 5 is a perfect square. ■

EXAMPLE G.23. The number 31, which ends in a 1, is equivalent to 1 mod 5. Using a computer system or checking the table in the back of the book (well, the number 30 isn't too large that you can't do this by hand) shows that $F_{30}=832,040$. And in accordance with the theorem we indeed do have that the number 31 divides this, in fact

$$832040=2^3\cdot 5\cdot 11\cdot 31\cdot 61.$$

G.3 ANOTHER LOOK AT $\mathbf{PSL}_2(\mathbb{F})$

Denote the set of lines in $\mathbb{F}^2$ as $\mathbb{FP}^1$. That is

$$\mathbb{FP}^1 \stackrel{\text{def}}{=} \mathbb{F}^2 \setminus \{\mathbf{0}\}\big/_{\asymp}.$$

We set up a dictionary on $\mathbb{FP}^1$ via $\begin{bmatrix} a \\ b \end{bmatrix} \rightsquigarrow \frac{a}{b}$ with something non-zero over zero being infinity. We never have to worry about $\frac{0}{0}$ since, as you recall, the zero vector is not allowed to lie in any line, in our definition. Notice that this dictionary is well defined since for all $r \neq 0$,

$$r \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} ra \\ rb \end{bmatrix} \rightsquigarrow \frac{ra}{rb} = \frac{a}{b} \leftsquigarrow \begin{bmatrix} a \\ b \end{bmatrix}.$$

One cool thing about this dictionary is that its obvious how to denote every line: if $v_2 \neq 0$, then the line

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \asymp \begin{bmatrix} v \\ 1 \end{bmatrix} \rightsquigarrow v$$

where $v = v_1 v_2^{-1}$. On the other hand if $v_2 = 0$, then $v_1 \neq 0$ and so

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ 0 \end{bmatrix} \asymp \begin{bmatrix} 1 \\ 0 \end{bmatrix} \rightsquigarrow \infty.$$

This means the $\|\mathbb{F}\| + 1$ lines in $\mathbb{F}^2$, i.e. the elements of $\mathbb{FP}^1$ are uniquely as either an element of $\mathbb{F}$ (accounting for $\|\mathbb{F}\|$ elements) or as ∞, giving rise the extra element corresponding to the $+1$ in the formula.

Another cool thing about this dictionary is that a matrix acts by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \bullet \frac{z_1}{z_2} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \bullet \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} az_1 + bz_2 \\ cz_1 + dz_2 \end{bmatrix} = \frac{az_1 + bz_2}{cz_1 + dz_2}.$$

If $z_2 \neq 0$, the vector

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \asymp \begin{bmatrix} z_1 z_2^{-1} \\ 1 \end{bmatrix} = \begin{bmatrix} z \\ 1 \end{bmatrix}$$

where $z := z_1 z_2^{-1}$. With this we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \bullet z = \frac{az + b}{cz + d}.$$

These are not functions, $\mathbb{F} \to \mathbb{F}$, but they are functions $\mathbb{F} \cup \{\infty\} \to \mathbb{F} \cup \{\infty\}$, and are known as *fractional linear transformations* (over $\mathbb{F}$). On the other hand, if $z_2 = 0$, then the point in question

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} z_1 \\ 0 \end{bmatrix} \asymp \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \infty.$$

Moreover, in this case we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \bullet \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix} = \frac{a}{c}.$$

This is also quite nice since if we were allowed to do calculus (which we aren't really!) and we took a limit of the function

$$\frac{az + b}{cz + d}$$

as $z \to \infty$, we'd get exactly this $\frac{a}{c}$. In this way, the statement that the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

thought of as an element of $\mathbf{PGL}_2(\mathbb{F})$ acts on $\mathbb{FP}^1$ identified with $\mathbb{F} \cup \{\infty\}$ via the fractional linear transformation

$$\frac{az+b}{cz+d}$$

shouldn't lead us to any confusion!

$$\text{Invertible Fractional Linear Transformations over } \mathbb{F} \overset{\text{1-1}}{\leftrightsquigarrow} \mathbf{PGL}_2(\mathbb{F})$$

and

$$\mathbb{FP}^1 \overset{\text{1-1}}{\leftrightsquigarrow} \mathbb{F} \cup \{\infty\}.$$

The last piece of the puzzle is: *How can we tell whether a fractional linear transformation comes from an element of* $\mathbf{SL}_2(\mathbb{F})$*?* Well, obviously, if we have the

$$\frac{az+b}{cz+d} \quad \text{with} \quad ad-bc=1,$$

then under our correspondence, certainly $\frac{az+b}{cz+d} \in \mathbf{PSL}_2(\mathbb{F})$. However, this determinant like quantity, $ad-bc$, is not well defined. For instance, looking at the fractional linear transformation $z \mapsto 4 \cdot z$, we'd think that this dude did not correspond to an element of $\mathbf{PSL}_2(\mathbb{F})$. However, on second glance we see that indeed the matrix (assuming characteristic of $\mathbb{F}$ is not 2)

$$\begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

has determinant 1, and so is in $\mathbf{SL}_2(\mathbb{F})$, and the corresponding fractional linear transformation is

$$\frac{2 \cdot z + 0}{0 \cdot z + \frac{1}{2}} = \frac{2 \cdot z}{\frac{1}{2}} = 4 \cdot z.$$

Therefore, our initial guess that $4 \cdot z \notin \mathbf{PSL}_2(\mathbb{F})$ is wrong.

Do we give up? NO! Like fools, we charge on! While the "determinant" of a fractional linear transformation is not well defined, we have for all $r \in F^\times$,

$$ad-bc \mapsfrom \frac{az+b}{cz+d} = \frac{arz+br}{crz+dr} \mapsto (ar)\cdot(dr)-(br)\cdot(cr) = (ad-bc)\cdot r^2.$$

Therefore, the determinant of a fractional linear transformation, albeit not well-defined, is well defined up to a perfect square, i.e., in fancier language, the map

$$\det\colon \mathbf{PGL}_2(\mathbb{F}) \to {}^{\mathbb{F}^\times}\!/_{\mathbb{F}^{\times 2}}$$

is well defined. To try to limit confusion, we call this map *the "determinant"* (with quotes). In less fancy language, while it does not make sense to ask what the determinant of a fractional linear transformation is, it does make sense to ask whether or not the determinant is a perfect square or not. [I]

Back to our original question: how can we tell if a fractional linear transformation is in $\mathbf{PSL}_2(\mathbb{F})$? Well, the "determinant" differs from 1 by a perfect square, i.e. it must be a perfect square. Conversely, if the "determinant" of

$$\frac{az+b}{cz+d}$$

is a perfect square, say r^2, then we can rewrite this very same fractional linear transformation can be written as

$$\frac{az+b}{cz+d} = \frac{\frac{a}{r}z+\frac{b}{r}}{\frac{c}{r}z+\frac{d}{r}}$$

which has "determinant" 1. Therefore, we have proven that

[I] Again, this is akin to defining the anti-derivative of a function up to a constant since every function has many anti-derivatives.

THEOREM G.24.

$$\mathbf{PSL}_2(\mathbb{F}) = \left\{ \frac{az+b}{cz+d} : ad - bc \in \mathbb{F}^{\times 2} \right\}.$$

■

COROLLARY G.25. *For all fields with at least one element which is not a perfect square, we have that* $\mathbf{PSL}_2(\mathbb{F})$ *has index 2 in* $\mathbf{PGL}_2(\mathbb{F})$.

REMARK G.26. The hypothesis of this corollary applies to all fields that we've dealt except $\mathbb{C}$ and $\mathbb{Z}_2$. With these exceptional fields, where every element is a perfect square, we have that $\mathbf{PSL}_2(\mathbb{F}) = \mathbf{PGL}_2(\mathbb{F})$.

PROOF. We've shown that $\mathbf{PSL}_2(\mathbb{F})$ is the kernel of the homomorphism

$$\det\colon \mathbf{PGL}_2(\mathbb{F}) \to {}^{\mathbb{F}^\times}\!/_{\mathbb{F}^{\times 2}},$$

and so long as there exists at least one non-perfect square in $\mathbb{F}$, we know that ${}^{\mathbb{F}^\times}\!/_{\mathbb{F}^{\times 2}} \cong \mathbb{Z}_2$, and so **KIT** applies. ■

THEOREM G.27. *The group* $\mathbf{PGL}_2(\mathbb{F})$ *acts sharply* 3*-transitively on* $\mathbb{F} \cup \{\infty\}$.

PROOF. The means for any triple of 3 distinct elements of $\mathbb{F} \cup \{\infty\}$ there exists a unique element of $\mathbf{PGL}_2(\mathbb{F})$ which maps it to any other triple. The easier triple to deal with is the triple $(0, \infty, 1)$. We need to show that any triple can be moved to $(0, \infty, 1)$ and that only the identity fixes 0, 1, and ∞.

(1) First we show that no non-identity fractional linear transformation can fix 0, ∞, and 1. Indeed, let

$$f(z) := \frac{az+b}{cz+d}$$

be an arbitrary fractional linear transformation. If $f(0) = 0$, then we must have that $b = 0$. If $f(\infty) = \infty$, then we must have that $c = 0$. Therefore, fractional linear transformation which fix both 0 and ∞ are of the form

$$f(z) := \frac{az}{d} = \overline{a}z$$

(where $\overline{a} = \frac{a}{d}$.) Lastly, if in addition we also have that $f(1) = 1$, this also means that $\overline{a} = 1$, and so our buddy $f(z)$ is forced to be $f(z) = z$, the identity fractional linear transformation, like I promised.

(2) Let (α, β, γ) be a triple of distinct points in $\mathbb{F} \cup \{\infty\}$. Then the fractional linear transformation

$$\frac{(z-\alpha)(\gamma-\beta)}{(z-\beta)(\gamma-\alpha)}$$

takes α to 0, β to ∞ and γ to 1, as we needed. (There is an extra step in checking this in the case that any of α, β, or γ happen to be ∞.) It's easy to see that this fractional linear transformation is invertible.

■

G.4 LOOKING AT $\mathbf{SL}_2(\mathbb{F})$ AND $\mathbf{PSL}_2(\mathbb{F})$ FOR MORE GENERAL FINITE FIELDS.

Many things carry over verbatim to the case that we look at finite fields $\mathbb{F}$ which aren't $\mathbb{Z}_p$.

We can either do the same exact calculation as we did in Chapter 11 or we can use Theorem G.27 to prove that for any finite field $\mathbb{F}_q$ we have

$$\mathbf{PGL}_2\left(\mathbb{F}_q\right) = (q-1)\cdot q \cdot (q+1) \quad \text{and} \quad \mathbf{PSL}_2\left(\mathbb{F}_q\right) = \frac{(q-1)\cdot q \cdot (q+1)}{2} \quad \text{and} \quad \mathbf{PSL}_2(\mathbb{F}_{2^m}) = \mathbf{PGL}_2(\mathbb{F}_{2^m})$$

where $\mathbb{F}_q$ is a (the) field of order q (necessarily a power of a prime). Now let's look at the two easiest special cases: $q = 4$ and $q = 9$.

G.4.1 $\mathbf{PSL}_2(\mathbb{F}_4) \cong \mathcal{A}_5$

In the case that $q = 4$, as a power of 2, we can't use a square root.

However, the polynomial $p(t) := t^2 + t + 1$ is irreducible in $\mathbb{Z}_2[t]$. The quotient

$$\mathbb{F}_4 \stackrel{\text{def}}{=} \mathbb{Z}_2[\omega] := \mathbb{Z}_2[t] \Big/ \langle\, p(t) \,\rangle$$

is therefore a field of order 4. (Where denoting the root of $t^2 + t + 1$, which exists in the bigger field, as ω). there are $4 + 1 = 5$ lines in $\mathbb{F}_4$. Right away, we should hope that

$$\mathbf{SL}_2(\mathbb{F}_4) \cong \mathcal{A}_5.$$

For starters, both have order

$$3 \cdot 4 \cdot 5 = 60,$$

so it would appear we're in luck. We just need to check that all elements of $\mathbf{SL}_2(\mathbb{F}_4)$ act as even permutations. Well, so far we know that the action, being faithful, (as $\mathbf{PSL}_2(\mathbb{F}) \circlearrowright \mathbb{FP}^1$ always is) gives us an injection

$$\mathbf{PSL}_2(\mathbb{F}_4) \to \mathcal{S}_5.$$

We know by the general theory that the only index 2 subgroup of $\mathcal{S}_5$ is $\mathcal{A}_5$ and so the image of this permutation representation must be $\mathcal{A}_5$ proving that

$$\mathbf{PSL}_2(\mathbb{F}_4) \cong \mathcal{A}_5.$$

Let's look at this in a bit more detail to help our intuition. Let's denote the five lines in $\mathbb{F}_4^2$ as we discussed in the previous section, i.e.

$$0 := \left\langle \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\rangle \qquad 1 \stackrel{\text{def}}{=} \left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\rangle$$

$$\omega := \left\langle \begin{bmatrix} \omega \\ 1 \end{bmatrix} \right\rangle \qquad 1 + \omega \stackrel{\text{def}}{=} \left\langle \begin{bmatrix} 1 + \omega \\ 1 \end{bmatrix} \right\rangle^{\text{I}}$$

$$\infty := \left\langle \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\rangle.$$

Let's check out how

$$\begin{pmatrix} 1 & \omega \\ 0 & 1 \end{pmatrix}$$

acts on $\mathbb{F}_4\mathbb{P}^1$. Well, by the last section, it's by the translation

$$z + \omega.$$

This sends 0 to ω and 1 to $\omega + 1$, i.e. it acts via the $\underline{2}\,\underline{2}$-cycle $(0\ \omega)(1\ \omega + 1)$. How about the matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

how does it act? Well, again, it's via the function

$$\frac{1}{z}$$

[I] Note that $\omega^2 + \omega + 1 = 0 \Rightarrow \omega^2 = -\omega - 1 = \omega + 1$.

which takes 0 to ∞, 1 to 1, and ω to $\omega+1$ (since $\omega\cdot(\omega+1)=\omega^2+\omega=\omega+\omega+1=1$, we have that $\frac{1}{\omega}=\omega+1$), and again this acts by a $\underline{2}\,\underline{2}$-cycle, this time it's $(0\ \infty)(\omega\ \omega+1)$. We should ask ourselves who acts like a $\underline{5}$-cycle, say $(0\ 1\ \omega\ \omega+1\ \infty)$. There's gotta be somebody since we're dealing with an isomorphism. Well, figuring out who is pretty easy as well.

Let's consider

$$\frac{az+b}{cz+d}$$

and solve for a, b, c, and d using the cycle we want, $(0\ 1\ \omega\ \omega+1\ \infty)$. Well, since $0\mapsto 1$ we have

$$\frac{a\cdot 0+b}{c\cdot 0+d}=\frac{b}{d}=1\Rightarrow b=d.$$

Next, since $\infty\mapsto 0$ we also have that

$$\frac{a}{d}=0\Rightarrow a=0.$$

Lastly, since $1\mapsto\omega$, we also have that

$$\frac{a+b}{c+d}=\omega\Rightarrow\frac{b}{c+b}=\omega\Rightarrow b=(c+b)\omega\Rightarrow(1\overset{\omega^2}{\overset{\shortparallel}{+}}\omega)\,b=c\omega\Rightarrow b\omega=c.$$

where we've made liberal use of the fact that $x=-x$ in $\mathbb{F}_4$ (or in any field of characteristic 2). Using all three of these facts shows us that

$$\frac{az+b}{cz+d}=\frac{b}{\omega\cdot bz+b}=\frac{1}{\omega\cdot z+1},$$

and we have determined our function! Let's check a couple of points. Plugging in ω shows us that

$$\omega\mapsto\frac{1}{\omega^2+1}=\frac{1}{\omega}=\omega+1,$$

which is what we needed.

G.4.2 $\mathbf{PSL}_2(\mathbb{F}_9)\cong\mathcal{A}_6$

This isomorphism will take a bit more effort to construct. We do it backwards from how to did the previous: we turn $\mathcal{A}_6$ into permutations in the way that we did in the previous section. Since there are 10 lines in $\mathbb{F}_9^2$, i.e. 10 elements in $\mathbb{F}_9\mathbb{P}^1$, and $10\neq 6$the isomorphism we need to find won't be nearly as obvious.

For starters, are there 10 things that $\mathcal{A}_6$ acts on in a fairly obvious way? Well, since no element can fix more than 2 things, $\underline{3}$ and $\underline{3}\,\underline{3}$–cycles must act as a $\underline{3}\,\underline{3}\,\underline{3}$-cycle. Moreover since $\underline{2}\,\underline{2}\,\underline{2}\,\underline{2}\,\underline{2}$-cycles are odd, a $\underline{2}\,\underline{2}$–cycle must act as a $\underline{2}\,\underline{2}\,\underline{2}\,\underline{2}$-cycle. A $\underline{5}$–cycle must act as a $\underline{5}\,\underline{5}$–cycle. Lastly, how must a $\underline{4}\,\underline{2}$–cycle act? Well, they must either act as $\underline{4}\,\underline{4}$–cycles, or as $\underline{4}\,\underline{2}\,\underline{2}\,\underline{2}$–cycles. All told, this means that only $\underline{3}$ and $\underline{3}\,\underline{3}$– cycles fix exactly one line.

Let's look at the action of

$$\mathcal{A}_6\circlearrowright\mathbf{Syl}_3\mathcal{A}_6.$$

Note that a Sylow 3-group has power $\eta_3(\frac{6\cdot5\cdot4\cdot3\cdot2\cdot1}{2})=2$ and has order $3^2=9$, and so

$$P_3:=\langle(1\ 2\ 3),(4\ 5\ 6)\rangle$$

is a Sylow 3-subgroup. Since it is generated by two disjoint $\underline{3}$–cycles, P_3 is isomorphic to $\mathbb{Z}_3\times\mathbb{Z}_3$. Recall to figure out how many elements there are in $\mathbf{Syl}_3\mathcal{A}_6$: we need to calculate the normalizer of P_3 and take its index in $\mathcal{A}_6$. As always, this normalizer includes the original group, P_3, but it also contains the $\underline{2}\,\underline{2}$-cycle

$$(1\ 2)(4\ 5).^{\text{II}}$$

[II] See Remark G.28.

Remark G.28. This $\underline{2}\,\underline{2}$–cycle is not obvious, in fact I missed it at first! Perhaps an easier, but less direct, way of figuring out the normalizer is doing the computation in $\mathcal{S}_6$ where it is obvious that the $\underline{2}\,\underline{2}\,\underline{2}$-cycle

$$(1\,4)(2\,5)(3\,6)$$

swaps the two generating $\underline{3}$–cycles of P_3 (which is also a Sylow 3-subgroup of $\mathcal{S}_6$), and, as such, normalizes it (it being P_3). Moreover, the part of the normalizer which fixes the two subgroups generated by the individual $\underline{3}$-cycles, $\langle (1\,2\,3) \rangle$ and $\langle (4\,5\,6) \rangle$, is clearly isomorphic to $\mathcal{S}_3 \times \mathcal{S}_3$. Therefore, $\mathrm{N}_{\mathcal{S}_6}(P_3)$ has order $6 \cdot 6 \cdot 2 = 72$ and index $\frac{720}{72} = 10$. Since the normalizer contains an odd permutation, (for instance $(1\,4)(2\,5)(3\,6)$ just mentioned) this is also the index of $\mathrm{N}_{\mathcal{A}_6}(P_3)$ in $\mathcal{A}_6$, see the footnote on the bottom of page 570.

This means, that the part of the normalizer of P_3 which fixes both $\mathbb{Z}_3$ factors of P_3 has order 18. Now, there's also

$$(1\,4\,2\,5)(3\,6)$$

which swaps the two $\mathbb{Z}_3$ factors generating P_3, giving rise to a total of 36 elements in the normalizer of P_3. This means there are a total of 10 Sylow 3-subgroups of $\mathcal{A}_6$ (pay dirt!)

We can actually make this a bit more concrete. There are a total of

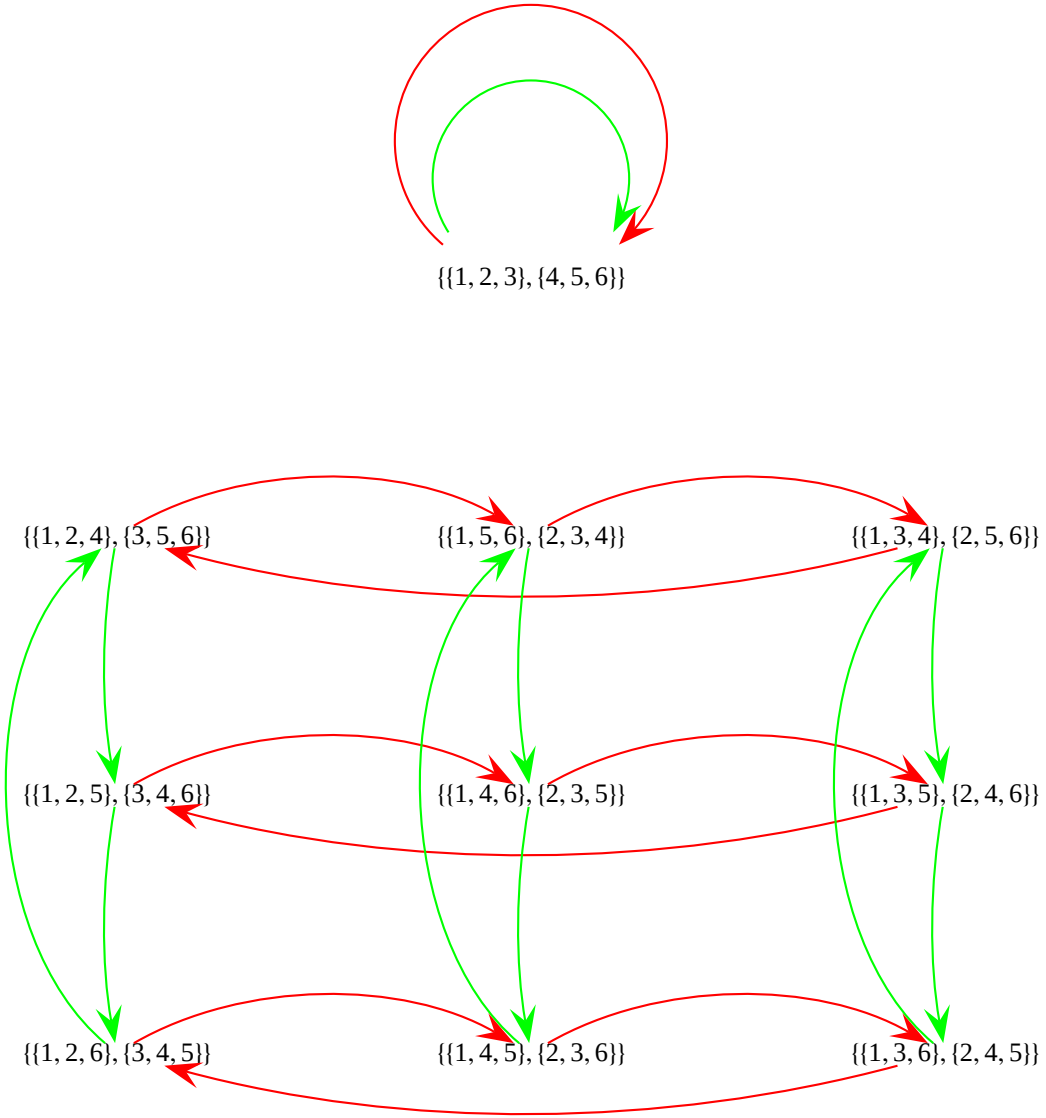
$$\frac{1}{2} \cdot \binom{6}{3} = \frac{1}{2} \cdot \frac{6!}{3! \cdot 3!} = \frac{1}{2} \cdot 5 \cdot 4 = 10$$

pairs of disjoint sets containing 3 elements in $\mathcal{A}_6$. For instance,

$$\{\{1, 2, 3\}, \{4, 5, 6\}\}$$

is one such pairing. Note that an element of $\sigma\mathcal{A}_6$ fixes this pair iff it is in the normalizer of P_3. By the strong form of **OST**, this means that the action on $\mathbf{Syl}_3\mathcal{A}_6$ and on these triples is the same.

Let's continue by examining what $(1\ 2\ 3)$ and $(4\ 5\ 6)$ do to these 10 guys: (The action of $(1\ 2\ 3)$ is in red, and the action of $(4\ 5\ 6)$ is in green.)



After they both fix $\{\{1, 2, 3\}, \{4, 5, 6\}\}$, they make a perfect grid! If we label this grid as follows:

$$\infty$$

$$0 \qquad 1 \qquad 2$$

$$i \qquad 1+i \qquad 2+i$$

$$2i \qquad 1+2i \qquad 2+2i,$$

then $(1\ 2\ 3)$ acts as the fractional linear transformation $f_1(z) := z + 1$, and $(4\ 5\ 6)$ acts as $f_2(z) := z + i$. Of course we cannot possibly done since these two guys do not generate $\mathcal{A}_6$, so we need to find the effect of at least one more element. Perhaps it's easier to look at this from the other point of view: turning a fractional linear transformation into a permutation. It shouldn't be difficult to convince yourself that, in general,

$$\mathbf{PSL}_2(\mathbb{F}) = \left\langle -\frac{1}{z}, \{a \in \mathbb{F} : x + a\} \right\rangle.$$

EXERCISE G.4. Prove this. That is prove, in general, that

$$\mathbf{PSL}_2(\mathbb{F}) = \left\langle -\frac{1}{z}, \{a \in \mathbb{F} : z + a\} \right\rangle.$$

SOLUTION

H Some Multiplication Tables

This example is 2^2 under the product operation.

$\mathbb{Z}_2 \times \mathbb{Z}_2$	$([0]_2, [0]_2)$	$([1]_2, [0]_2)$	$([0]_2, [1]_2)$	$([1]_2, [1]_2)$
$([0]_2, [0]_2)$	$([0]_2, [0]_2)$	$([1]_2, [0]_2)$	$([0]_2, [1]_2)$	$([1]_2, [1]_2)$
$([1]_2, [0]_2)$	$([1]_2, [0]_2)$	$([0]_2, [0]_2)$	$([1]_2, [1]_2)$	$([0]_2, [1]_2)$
$([0]_2, [1]_2)$	$([0]_2, [1]_2)$	$([1]_2, [1]_2)$	$([0]_2, [0]_2)$	$([1]_2, [0]_2)$
$([1]_2, [1]_2)$	$([1]_2, [1]_2)$	$([0]_2, [1]_2)$	$([1]_2, [0]_2)$	$([0]_2, [0]_2)$

This example is $\mathcal{S}_3$, the symmetric group on three letters, under composition.

$\mathcal{S}_2$	(1)	$(1\,2\,3)$	$(1\,3\,2)$	$(1\,2)$	$(2\,3)$	$(1\,3)$
(1)	(1)	$(1\,2\,3)$	$(1\,3\,2)$	$(1\,2)$	$(2\,3)$	$(1\,3)$
$(1\,2\,3)$	$(1\,2\,3)$	$(1\,3\,2)$	$(1\,3\,2)$	$(1\,3)$	$(1\,2)$	$(2\,3)$
$(1\,3\,2)$	$(1\,3\,2)$	(1)	$(1\,3\,2)$	$(2\,3)$	$(1\,3)$	$(1\,3)$
$(1\,2)$	$(1\,2)$	$(2\,3)$	$(1\,3)$	(1)	$(1\,2\,3)$	$(1\,3\,2)$
$(2\,3)$	$(2\,3)$	$(1\,3)$	$(1\,2)$	$(1\,3\,2)$	(1)	$(1\,2\,3)$
$(1\,3)$	$(1\,3)$	$(1\,2)$	$(2\,3)$	$(1\,2\,3)$	$(1\,3\,2)$	(1)

This example is $\mathcal{D}_4$, the symmetries of a square under composition.

$\mathcal{D}_4$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$
$\mathbf{1}$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$
ρ	ρ	ρ^2	ρ^3	$\mathbf{1}$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$
ρ^2	ρ^2	ρ^3	$\mathbf{1}$	ρ	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$
ρ^3	ρ^3	$\mathbf{1}$	ρ	ρ^2	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$
$\mathbf{r}$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{1}$	ρ	ρ^2	ρ^3
$\mathbf{r}\rho$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	ρ^3	$\mathbf{1}$	ρ	ρ^2
$\mathbf{r}\rho^2$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	ρ^2	ρ^3	$\mathbf{1}$	ρ
$\mathbf{r}\rho^3$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	ρ	ρ^2	ρ^3	$\mathbf{1}$

This example is Q_8 under quaternion multiplication.

Q_8	1	$\mathbf{i}$	-1	$-\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$	$-\mathbf{j}$	$-\mathbf{k}$
1	1	$\mathbf{i}$	-1	$-\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$	$-\mathbf{j}$	$-\mathbf{k}$
$\mathbf{i}$	$\mathbf{i}$	-1	$-\mathbf{i}$	1	$\mathbf{k}$	$-\mathbf{j}$	$-\mathbf{k}$	$\mathbf{j}$
-1	-1	$-\mathbf{i}$	1	$\mathbf{i}$	$-\mathbf{j}$	$-\mathbf{k}$	$\mathbf{j}$	$\mathbf{k}$
$-\mathbf{i}$	$-\mathbf{i}$	1	$\mathbf{i}$	-1	$-\mathbf{k}$	$\mathbf{j}$	$\mathbf{k}$	$-\mathbf{j}$
$\mathbf{j}$	$\mathbf{j}$	$-\mathbf{k}$	$-\mathbf{j}$	$\mathbf{k}$	-1	$\mathbf{i}$	1	$-\mathbf{i}$
$\mathbf{k}$	$\mathbf{k}$	$\mathbf{j}$	$-\mathbf{k}$	$-\mathbf{j}$	$-\mathbf{i}$	-1	$\mathbf{i}$	1
$-\mathbf{j}$	$-\mathbf{j}$	$\mathbf{k}$	$\mathbf{j}$	$-\mathbf{k}$	1	$-\mathbf{i}$	-1	$\mathbf{i}$
$-\mathbf{k}$	$-\mathbf{k}$	$-\mathbf{j}$	$\mathbf{k}$	$\mathbf{j}$	$\mathbf{i}$	1	$-\mathbf{i}$	-1

This example is $\mathcal{A}_4$ with

$1 = (1)$ $2 = (1\,2)(3\,4)$ $3 = (1\,3)(2\,4)$ $4 = (1\,4)(2\,3)$ $5 = (1\,2\,3)$ $6 = (1\,3\,4)$
$7 = (2\,4\,3)$ $8 = (1\,4\,2)$ $9 = (1\,3\,2)$ $10 = (2\,3\,4)$ $11 = (1\,2\,4)$ $12 = (1\,4\,3)$.

$\mathcal{A}_4$	1	2	3	4	5	6	7	8	9	10	11	12
1	1	2	3	4	5	6	7	8	9	10	11	12
2	2	1	4	3	7	8	5	6	12	11	10	9
3	3	4	1	2	8	7	6	5	10	9	12	11
4	4	3	2	1	6	5	8	7	11	12	9	10
5	5	6	7	8	9	10	11	12	1	2	3	4
6	6	5	8	7	11	12	9	10	4	3	2	1
7	7	8	5	6	12	11	10	9	2	1	4	3
8	8	7	6	5	10	9	12	11	3	4	1	2
9	9	10	11	12	1	2	3	4	5	6	7	8
10	10	9	12	11	3	4	1	2	8	7	6	5
11	11	12	9	10	4	3	2	1	6	5	8	7
12	12	11	10	9	2	1	4	3	7	8	5	6

In the following table the notation ±±±± (each ± potentially different) is short-hand for the quaternion

$$\pm\frac{1}{2} \pm \frac{1}{2}\mathbf{i} \pm \frac{1}{2}\mathbf{j} \pm \frac{1}{2}\mathbf{k}.$$

For instance, +−−+ is shorthand for the quaternion

$$+\frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}$$

and −−−+ is shorthand for the quaternion

$$-\frac{1}{2} - \frac{1}{2}\mathbf{i} - \frac{1}{2}\mathbf{j} + \frac{1}{2}\mathbf{k}.$$

Of course this notation is used here only as a space-saving device!

Recall that the 24 elements of $B(\blacktriangleright^3)$ consist of the 16 quaternions of that form along with the 8 quaternions in Q_8.

$B\left(\triangleright^{3}\right)$	1	**i**	−1	**−i**	**j**	**k**	**−j**	**−k**	+ + + +	− + − −	− − − −	+ − + −
1	1	**i**	−1	**−i**	**j**	**k**	**−j**	**−k**	+ + + +	− + − −	− − − −	+ − + −
i	**i**	−1	**−i**	1	**k**	**−j**	**k**	**j**	− + − +	− − − −	+ − + −	+ + + +
−1	−1	**−i**	1	**i**	**−j**	**−k**	**j**	**−k**	− − − −	+ − + −	+ + + +	− + − +
−i	**−i**	1	**i**	−1	**−k**	**j**	**k**	**−j**	+ − + −	− + − +	− − − −	+ − + −
j	**j**	**−k**	**−j**	**k**	−1	**i**	1	**−i**	− + + −	+ + − −	+ − − +	− − + +
k	**k**	**j**	**−k**	**−j**	**−i**	−1	**i**	1	− − + +	− + + −	+ + − −	+ − − +
−j	**−j**	**k**	**j**	**−k**	1	**−i**	−1	**i**	+ − − +	− − + +	− + + −	+ + − −
−k	**−k**	**−j**	**k**	**j**	**i**	1	**−i**	−1	9	+ − − +	− + + −	− + + −
+ + + +	+ + + +	− + + −	− − − −	+ − − +	− − + +	− + − +	+ + − −	+ − + −	− + + +	− + − +	− − − −	+ − + −
− + − +	− + −+	− − + +	+ − + −	+ + − −	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
− − − −	− − −−	+ − − +	+ + + +	− + + −	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
+ − + −	+ − + −	+ + − −	− + − +	− − + +	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
− + + −	+ − + −	+ − − −	+ − − +	− + + +	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
− − + +	− − ++	+ − + −	+ + − −	− + − +	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
+ − − +	+ − − +	+ + + +	− + + −	− − − −	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
+ + − −	+ + − −	− + − +	− − + +	+ − + −	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
+ − − −	+ − − −−	+ + − +	− − − +	− − + −	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
+ + + −	+ + + −	− + − −	− − − +	+ − + +	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
− + + +	− + ++	− − + −	+ − − −	+ + − +	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
− − − +	− − −+	+ − + +	+ + + −	− + − −	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
+ − + +	+ − + +	+ + + −	− + − −	− − − +	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
+ + − +	+ + − +	− + + +	− − + −	+ − − −	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
− + − −	− + −−	− − − +	+ − + +	+ + + −	5	6	7	**−k**	9	− + − +	− − − −	+ − + −
− − + −	− − +−	+ − − +	+ + − +	− + + −	5	6	7	**−k**	9	− + − +	− − − −	+ − + −

$B\left(\blacktriangleright^3\right)$	− + + −	− − + +	+ − − +	+ + − −	+ − − −	+ + + −	− + + +	− − − +	+ − + +	+ + − +	− + − −	− − + −
1	− + + −	− − + +	+ − − +	+ + − −	+ − − −	+ + + −	− + + +	− − − +	+ − + +	+ + − +	− + − −	− − + −
i	− − + +	+ − − +	+ + − −	− + + −	+ + + −	− + + +	− − − +	+ − − −	+ + − +	− + − −	− − + −	− − − +
−1	+ − − +	+ + − −	− + +−	− − ++	− − −+	− − −+	+ − −−	+ + +−	− + −−	− − +−	+ − + +	+ + − +
−i	+ + − −	− + + −	− − + +	+ − − +	− − − +	+ − − −	+ + + −	− + + +	− − + −	+ − + +	+ + − +	+ + + −
j	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
k	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
−j	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
−k	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
+ + ++	1	**i**	−1	4	1	6	−1	**−k**	9	− + −+	− − −−	+ − + −
− + −+	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
− − −−	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
+ − +−	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
− + +−	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
− − ++	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
+ − −+	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
+ + −−	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
+ − −−	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
+ + +−	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
− + ++	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
− − −+	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
+ − ++	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
+ + −+	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
− + −−	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −
− − +−	1	**i**	−1	4	5	6	7	**−k**	9	− + −+	− − −−	+ − + −

Here are the most common groups which show up in the text along with their order.

Group	$\mathbb{Z}_n$	$\mathcal{S}_n$	$\mathbf{SL}_2\left(\mathbb{Z}_p\right)$	$\mathcal{U}(\mathbb{Z}_n)$	$\mathbb{H}^\times$	$\mathbb{S}^3$	$\mathbb{C}^\times$	$\mathbb{S}^1$	$\mathbb{Z}$	$\mathcal{D}_n$	$B\mathcal{D}_n$
Order	n	$n!$	(p^3-p)	$\vartheta(n)$	∞	∞	∞	∞	∞	$2n$	$4n$
Abelian(?)	yes	no	no	yes	no	no	yes	yes	yes	no	no

Group	$\mathcal{A}_n$	$\{\{n\}\}$	$\mathbf{GL}_2\left(\mathbb{Z}_p\right)$	$B\left(\blacktriangleright^3\right)$	$B\left(\blacklozenge^3\right)$	$B\left(\bigstar^3\right)$	$S\left(\blacktriangleright^3\right)$	$S\left(\blacklozenge^3\right)$	$S\left(\bigstar^3\right)$
Order	$\frac{n!}{2}$	n	$(p-1)(p^3-p)$	24	48	120	12	24	60
Abelian(?)	no	yes	no	no	no	no	no	no	no

I Number of Groups of Given Order

The second row of each segment denotes the number of non–isomorphic groups of order given by the top row of the segment. The last row gives the number of abelian groups of that same order. This latter number, the number of abelian group of order n, according to Corollary 13.55 is given by

$$\prod_{p\geq 0,\text{ prime}} \pi(\nu_p(n)).$$

Order	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Tot #	1	1	1	2	1	2	1	5	2	2	1	5	1	2	1
Abel	1	1	1	2	1	1	1	3	2	1	1	2	1	1	1

Order	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Tot#	14	1	5	1	5	2	2	1	15	2	2	5	4	1	4
Abel	5	1	2	1	2	1	1	1	3	2	1	3	2	1	1

Order	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
Tot #	1	51	1	2	1	14	1	2	2	14	1	6	1	4	2
Abel	1	7	1	1	1	4	1	1	1	3	1	1	1	2	2

Order	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Tot #	2	1	52	2	5	1	5	1	15	2	13	2	2	1	13
Abel	1	1	5	2	2	1	2	1	3	1	3	1	1	1	2

Order	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
Tot #	1	2	4	267	1	4	1	5	1	4	1	50	1	2	3
Abel	1	1	2	11	1	1	1	1	1	1	1	6	1	1	2

Order	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
Tot #	4	1	6	1	52	15	2	1	15	1	2	1	12	1	10
Abel	2	1	1	1	5	5	1	1	2	1	1	1	3	1	2

Order	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105
Tot #	1	4	2	2	1	231	1	5	2	16	1	4	1	14	2
Abel	1	2	1	1	1	7	1	2	2	4	1	1	1	3	1

Order	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
Tot #	2	1	45	1	6	2	43	1	6	1	5	4	2	1	47
Abel	1	1	6	1	1	1	5	1	1	1	2	2	1	1	3

Order	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135

Tot #	2	2	1	4	5	16	1	2328	2	4	1	10	1	2	5
Abel	2	1	1	2	3	2	1	15	1	1	1	2	1	1	3
Order	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150
Tot#	15	1	4	1	11	1	2	1	197	1	2	6	5	1	13
Abel	3	1	1	1	2	1	1	1	10	1	1	2	2	1	2
Order	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165
Tot#	1	12	2	4	2	18	1	2	1	238	1	55	1	5	2
Abel	1	3	2	1	1	2	1	1	1	7	1	5	1	2	1
Order	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180
Tot#	2	1	57	2	4	5	4	1	4	2	42	1	2	1	37
Abel	1	1	3	2	1	2	2	1	1	2	5	1	1	1	4
Order	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195
Tot#	1	4	2	12	1	6	1	4	13	4	1	1543	1	2	2
Abel	1	1	1	3	1	1	1	2	3	1	1	11	1	1	1
Order	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210
Tot#	12	1	10	1	52	2	2	2	12	2	2	2	51	1	12
Abel	4	1	2	1	6	1	1	1	2	1	1	2	5	1	1
Order	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225
Tot#	1	5	1	2	1	177	1	2	2	15	1	6	1	197	6
Abel	1	2	1	1	1	9	1	1	1	2	1	1	1	7	4
Order	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240
Tot#	2	1	15	1	4	2	14	1	16	1	4	2	4	1	208
Abel	1	1	2	1	1	1	3	1	2	1	2	1	1	1	5
Order	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255
Tot#	1	5	67	5	2	4	1	12	1	15	1	46	2	2	1

Abel	1	2	7	2	2	1	1	3	1	3	1	4	1	1	1
Order	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270
Tot#	56092	1	6	1	15	2	2	1	39	1	4	1	4	1	30
Abel	22	1	1	1	2	2	1	1	3	1	1	1	2	1	3
Order	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285
Tot#	1	54	5	2	4	10	1	2	4	40	1	4	1	4	2
Abel	1	5	1	1	2	2	1	1	2	3	1	1	1	2	1
Order	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300
Tot#	4	1	1045	2	4	2	5	1	23	1	14	5	2	1	49
Abel	1	1	14	2	1	1	2	1	2	1	3	3	1	1	4
Order	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315
Tot#	2	2	1	42	2	10	1	9	2	6	1	61	1	2	4
Abel	1	1	1	5	1	2	1	2	1	1	1	3	1	1	2
Order	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330
Tot#	4	1	4	1	1640	1	4	1	176	2	2	2	15	1	12
Abel	2	1	1	1	11	1	1	1	10	2	1	1	3	1	1
Order	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345
Tot#	1	4	5	2	1	228	1	5	1	15	1	18	5	12	1
Abel	1	2	2	1	1	5	1	2	1	2	1	2	3	3	1
Order	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360
Tot#	2	1	12	1	10	14	195	1	4	2	5	2	2	1	162
Abel	1	1	2	1	2	3	7	1	1	1	2	1	1	1	6
Order	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375
Tot#	2	2	3	11	1	6	1	42	2	4	1	15	1	4	7
Abel	2	1	2	2	1	1	1	5	2	1	1	2	1	1	3

Order	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390
Tot#	12	1	60	1	11	2	2	1	20169	2	2	4	5	1	12
Abel	3	1	3	1	2	1	1	1	15	1	1	2	2	1	1

Order	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405
Tot#	1	44	1	2	1	30	1	2	5	221	1	6	1	5	16
Abel	1	6	1	1	1	4	1	1	1	10	1	1	1	2	5

Order	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420
Tot#	6	1	46	1	6	1	4	1	10	1	235	2	4	1	41
Abel	1	1	3	1	1	1	2	1	2	1	7	1	1	1	2

Order	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435
Tot#	1	2	2	14	2	4	1	4	2	4	1	775	1	4	1
Abel	1	1	2	3	2	1	1	2	1	1	1	15	1	1	1

Order	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450
Tot#	5	1	6	1	51	13	4	1	18	1	2	1	1396	1	34
Abel	2	1	1	1	3	4	1	1	2	1	1	1	11	1	4

This table lists the number of groups up to isomorphism of a given order, i.e. the number of "distinct" groups of a given order where I consider isomorphic groups not distinct. As we go along in this book, you should verify that the theorems we are proving are true by checking this table. If nothing else, this should help you gain confidence. Now I'm going to make a table of the number of p-groups there are of a given order. It is rather interesting that most groups are really p-groups. For instance, there are a Totof $49,910,529,484$ non-isomorphic groups of order less than or equal to 2000. However, a whopping $49,487,365,422$ of them have order $2^{10} = 1,024$. Therefore there are only a mere $423,164,062$ groups of order less than 2000 which do not have order 1,024! (49 billion versus 423 million, less than 1 percent!!)

In the table, red numbers are prime numbers. Notice for each such number there is only one group (namely the cyclic group of that order) of that order, which is in accordance with Corollary 10.69. Numbers in yellow are twice a prime. Notice in each of these cases there are two groups (one is isomorphic to $\mathbb{Z}_{2p}$ and the other is isomorphic to $\mathcal{D}_{2p}$) which is in accordance with Theorem 10.134. Numbers in blue are the square of a prime number. Notice in each of these cases there are two groups of that order. This is in accordance with Theorem **??**. (The number 4 is green since it is colored both yellow and blue... ain't that cute?)

Order	2^1	2^2	2^3	2^4	2^5	2^6	2^7	2^8	2^9	2^{10}
Tot #	1	2	5	14	51	267	2,328	56,092	10,494,213	49,487,365,422
Abel	1	2	3	5	7	11	15	22	30	42

Order	3^1	3^2	3^3	3^4	3^5	3^6	3^7
Tot #	1	2	5	15	67	504	9,310
Abel	1	2	3	5	7	11	15

Order	5^1	5^2	5^3	5^4	5^5	5^6	5^7
Tot	1	2	5	15	77	684	34,297
Abel	1	2	3	5	7	11	15

Order	7^1	7^2	7^3	7^4	7^5	7^6	7^7
Tot #	1	2	5	15	83	860	113,147
Abel	1	2	3	5	7	11	15

Order	11^1	11^2	11^3	11^4	11^5	11^6
Tot #	1	2	5	15	87	1192
Abel	1	2	3	5	7	11

Order	13^1	13^2	13^3	13^4	13^5	13^6
Tot #	1	2	5	15	97	1476
Abel	1	2	3	5	7	11

Table I.2: Number of p–groups (up to isomorphism)

J Groups of Order 16

Without proof, the following represent all 14 groups of order 16, up to isomorphism (of course). This appendix can be understood after reading Chapter 12, Semi-Direct Products.

$$
\begin{aligned}
G_1 &= \mathbb{Z}_{16} \\
G_2 &= \mathbb{Z}_4 \times \mathbb{Z}_4 \\
G_3 &= \mathbb{Z}_2^2 \rtimes_{\mathbf{A}} \mathbb{Z}_4 \qquad (\text{With } \mathbf{A} = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)) \\
G_4 &= \mathbb{Z}_4 \rtimes_{-1} \mathbb{Z}_4 \\
G_5 &= \mathbb{Z}_8 \times \mathbb{Z}_2 \\
G_6 &= \mathbb{Z}_8 \rtimes_5 \mathbb{Z}_2 \\
G_7 &= \mathcal{D}_8 \; ; \; (= \mathbb{Z}_8 \rtimes_{-1} \mathbb{Z}_2) \\
G_8 &= \mathbb{Z}_8 \rtimes_3 \mathbb{Z}_2 \\
G_9 &= B\mathcal{D}_4 \\
G_{10} &= \mathbb{Z}_4 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \\
G_{11} &= \mathcal{D}_4 \times \mathbb{Z}_2 \\
G_{12} &= \mathcal{Q}_8 \times \mathbb{Z}_2 \\
G_{13} &= (\mathbb{Z}_4 \times \mathbb{Z}_2) \langle x, y \rangle \rtimes \mathbb{Z}_2 \langle z \rangle \quad (\text{With } zxz^{-1} = x \quad \text{and} \quad zyz^{-1} = x^2 y) \\
G_{14} &= \mathbb{Z}_2^4.
\end{aligned}
$$

In my opinion, the strangest of these 14 groups is the second to last, G_{13}. This group has three generators: x which has order 4, y which has order 2, and z which also has order 2. The element x is central while the element z conjugates y to x^2y. Note that this makes sense this latter product, like y, has order 2. Interestingly, it is the only group of order 16 with center cyclic of order 4.

J.1 Data Tables

Here is some data regarding these groups.

1 of 14 $(\mathbb{Z}_{16})$

(a) $\mathcal{C}(\mathbb{Z}_{16}) = \mathbb{Z}_{16}$

(b) Order table is

$\mathbb{Z}_8 \times \mathbb{Z}_2$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	1	2	4	8

2 of 14 $(\mathbb{Z}_4 \times \mathbb{Z}_4)$

(a) $\mathcal{C}(\mathbb{Z}_4 \times \mathbb{Z}_4) = \mathbb{Z}_4 \times \mathbb{Z}_4$

(b) Order table is

$\mathbb{Z}_8 \times \mathbb{Z}_2$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	3	12	0	0

3 of 14 (G_3)

(a) $\mathcal{C}(G_3) = \left\langle xy, z^2 \right\rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$

(b) Order table is

$\mathbf{G}_3$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	7	6	8	0

‘

4 of 14 (G_4)

(a) $\mathcal{C}(G_4) = \left\langle x^2, y^2 \right\rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$

(b) Order table is

$\mathbf{G}_4$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	3	12	0	0

5 of 14 $(\mathbb{Z}_8 \times \mathbb{Z}_2)$

(a) $\mathcal{C}(\mathbb{Z}_8 \times \mathbb{Z}_2) = \mathbb{Z}_8 \times \mathbb{Z}_2$

(b) Order table is

$\mathbb{Z}_8 \times \mathbb{Z}_2$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	3	4	8	0

6 of 14 (G_6)

(a) $\mathcal{C}(G_6) = \left\langle x^2 \right\rangle \cong \mathbb{Z}_4$

(b) Order table is

$\mathbf{G}_6$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	3	4	8	0

7 of 14 $(\mathcal{D}_8)$

(a) $\mathcal{C}(\mathcal{D}_8) = \left\langle \rho^4 \right\rangle \cong \mathbb{Z}_2$.

(b) Order table is

$\mathcal{D}_8$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	9	2	4	0

8 of 14 (G_8)

(a) $\mathcal{C}(G_8) = \{e, x^4\} \cong \mathbb{Z}_2$.

(b) Order table is

$\mathbf{G}_8$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	5	6	4	0

9 of 14 $(B\mathcal{D}_4)$

(a) $\mathcal{C}(B\mathcal{D}_4) = \{\pm 1\} \cong \mathbb{Z}_2$.

(b) Order table is

$\mathbf{B}\mathcal{D}_4$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	1	10	4	0

10 of 14 $(\mathbb{Z}_2^2 \times \mathbb{Z}_4)$

(a) $\mathcal{C}(\mathbb{Z}_2^2 \times \mathbb{Z}_4) = (\mathbb{Z}_2^2 \times \mathbb{Z}_4)$

(b) Order table is

$\mathbb{Z}_2{}^2 \times \mathbb{Z}_4$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	7	8	0	0

11 of 14 $(\mathcal{D}_4 \times \mathbb{Z}_2)$

(a) $\mathcal{C}(\mathcal{D}_4 \times \mathbb{Z}_2) = \left\langle (\rho^2, [0]_2), (\mathbf{1}, [1]_2) \right\rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$

(b) Order table is

$\mathcal{D}_4 \times \mathbb{Z}_2$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	11	4	0	0

12 of 14 $(\mathcal{Q}_8 \times \mathbb{Z}_2)$

(a) $\mathcal{C}(Q_8 \times \mathbb{Z}_2) = \langle\, (-1, [\,0\,]_2), (1, [\,1\,]_2) \,\rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$

(b) $[Q_8 \times \mathbb{Z}_2, Q_8 \times \mathbb{Z}_2] = \langle\, (-1, [\,0\,]_2) \,\rangle \cong \mathbb{Z}_2$

(c) $(Q_8 \times \mathbb{Z}_2)_{ab} \cong \mathbb{Z}_2^3$

(d) Order table is

$Q_8 \times \mathbb{Z}_2$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	3	12	0	0

13 of 14 (G_{13})

(a) $\mathcal{C}(G_{13}) = \langle\, x \,\rangle \cong \mathbb{Z}_4$

(b) $[G_{13}, G_{13}] = \left\langle\, x^2 \,\right\rangle \cong \mathbb{Z}_2$

(c) $(G_{13})_{ab} \cong \mathbb{Z}_2^3$

(d) Order table is

G_{13}	n	1	2	4	8	16
	$\mathbf{o}_n$	1	7	8	0	0

14 of 14 $(\mathbb{Z}_2^4)$

(a) $\mathcal{C}(\mathbb{Z}_2^4) = \mathbb{Z}_2^4$

(b) $\left[\mathbb{Z}_2^4, \mathbb{Z}_2^4\right] = \mathfrak{e}$

(c) $\left(\mathbb{Z}_2^4\right)_{ab} = \mathbb{Z}_2^4$

(d) Order table is

$\mathbb{Z}_2{}^4$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	15	0	0	0

J.2 Why are they different

Evidently, there are only two pairs of order tables which are the same. They are the tables for G_4 and $Q_8 \times \mathbb{Z}_2$ and the tables for $\mathbb{Z}_8 \times \mathbb{Z}_2$ and G_6. Clearly the latter pair of groups are not isomorphic. This is because $\mathbb{Z}_8 \times \mathbb{Z}_2$ is abelian, while G_6 is not. However proving that $G_4 \ncong Q_8 \times \mathbb{Z}_2$ presents a more difficult problem![I]

To make matters worse, these two groups, in addition to having identical order tables, have isomorphic centers. This renders our two most powerful methods of proving groups are not isomorphic useless. To show they are non-isomorphic, we prove that the abelianization of G_4 is $\mathbb{Z}_4 \times \mathbb{Z}_2$ and the abelianization of $Q_8 \times \mathbb{Z}_2$ is isomorphic to $\mathbb{Z}_2^2$. Indeed, if we mod the first group out by $\left\langle x^2 \right\rangle$, we see that the quotient is isomorphic to $\mathbb{Z}_4 \times \mathbb{Z}_2$, but if we mod out the second by $\langle (-1, [0]_2) \rangle$ we see the quotient is isomorphic to $\mathbb{Z}_2^3$.

In the following exercises, you may assume, without proof, that (up to isomorphism) these are the only groups of order 16.

J.3 Exercises

J.3.1 Exploring these groups

Exercise J.1.

SOLUTION

(1) Verify that G_{13} has the stated center.

(2) Verify that G_{13} has the stated commutator.

(3) Verify that G_{13} has the stated abelianization.

(4) Verify that the only perfect square in G_{13}, besides the identity, is x^2.
Hint! You can do this quickly by using the fact that $G_{13}/\left\langle x^2 \right\rangle \cong \mathbb{Z}_2^2$.

(5) Verify the order table for G_{13}.

J.3.2 Classification Exercises

Exercise J.2. This exercise can be done after reading the chapter on quotient groups. Let G be the following group

SOLUTION

$$\left\langle (1\,2\,3\,4)\overset{\sigma}{=}(5\,6\,7\,8), (1\,5)(2\,6)\overset{\tau_1}{=}(3\,7)(4\,8), (5\,7)\overset{\tau_2}{=}(6\,8) \right\rangle.$$

Show that G has order 16. Which group on the table above is it?

[I] We come across G_4 in Example 11.11 on page 604.

Exercise J.3. Consider $\pi\colon \mathcal{D}_4 \to {}^{\mathcal{D}_4}\!/_{\langle \rho^2 \rangle}$, the standard quotient map. Define G as the set

$$\{\,(\alpha_1, \alpha_2) \in \mathcal{D}_4 \times \mathcal{D}_4 : \pi(\alpha_1) = \pi(\alpha_2)\,\}.$$

(1) Prove that G is a a subgroup of $\mathcal{D}_4 \times \mathcal{D}_4$.

(2) Prove that $|\,G\,| = 16$.

(3) Which of the groups of order 16 must G be isomorphic to?

Solution

Exercise J.4. Consider $\pi\colon Q_8 \to {}^{Q_8}\!/_{\langle -1 \rangle}$, the standard quotient map. Define G as

$$\{\,(\mathsf{q}, \mathsf{p}) \in Q_8 \times Q_8 : \pi(\mathsf{q}) = \pi(\mathsf{p})\,\}.$$

(1) Prove that G is a a subgroup of $Q_8 \times Q_8$.

(2) Prove that $|\,G\,| = 16$.

(3) Which of the groups of order 16 must G be isomorphic to?

Solution

Exercise J.5. Define G as

$$\{(\alpha, \mathsf{z}) \in \mathcal{D}_4 \times \{\{4\}\} : \det \alpha = \mathsf{z}^2\}.$$

(1) Prove that G is a a subgroup of $\mathcal{D}_4 \times \{\{4\}\}$.

(2) Prove that $|\,G\,| = 16$.

(3) Which of the groups of order 16 must G be isomorphic to?

Hint! Unlike when we did this construction for with $\mathcal{D}_3$ replacing $\mathcal{D}_4$ we ended up with $B\mathcal{D}_3$, the answer to this question is **not** $B\mathcal{D}_4$.

Solution

Exercise J.6. Define a homomorphism $\pi\colon Q_8 \to \mathbb{Z}_2$ via $\pi(q) = [\,0\,]_2$ if $\mathsf{q} \in \langle\,\mathbf{i}\,\rangle$ and $\pi(\mathsf{q}) = [\,1\,]_2$ otherwise. (It is the composition of the quotient map,

$$Q_8 \to {}^{Q_8}\!/_{\langle\,\mathbf{i}\,\rangle} \xrightarrow{\cong} \mathbb{Z}_2$$

Solution

where the first map is the quotient and the second is the unique isomorphism of the two element group which is the quotient with $\mathbb{Z}_2$.) Define G as

$$\{(\mathsf{q}, [\, a \,]_4) \in Q_8 \times \mathbb{Z}_4 : \pi(\mathsf{q}) = [\, a \,]_2\}.$$

(1) Prove that G is a a subgroup of $\mathcal{D}_4 \times \{\{4\}\}$.

(2) Prove that $|\, G \,| = 16$.

(3) Which of the listed groups of order 16 must G be isomorphic to?

Exercise J.7. Consider the quotient map

$$\pi \colon \mathcal{D}_4 \to \mathcal{D}_4 \big/ \langle \rho^2 \rangle. \quad (\cong \mathbb{Z}_2 \times \mathbb{Z}_2)$$

Solution

Define G as the group

$$\{(w, w') \in \mathcal{D}_4 \times \mathcal{D}_4 : \pi(w) = \pi(w')\}.$$

(1) Prove that G is a a subgroup of $\mathcal{D}_4 \times \mathcal{D}_4$.

(2) Prove that $|\, G \,| = 16$.

(3) Which of the listed groups of order 16 must G be isomorphic to?

Exercise J.8. Let

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 := \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$ [1]

Solution

and define G as the subgroup of $\mathbf{GL}_2(\mathbb{C})$ generated by these three matrices.

(1) Prove that $|\, G \,| = 16$.

(2) Which of the groups of order 16 must G be isomorphic to?

Exercise J.9. (**For those who have read Chapter 13, Sylow's Theorems**). Since

$$|\, \mathbf{GL}_2(\mathbb{Z}_3) \,| = 3 \cdot (3^2 - 1)\,(3 - 1) = 3 \cdot 2^4$$

Solution

the Sylow 2–subgroup of $\mathbf{GL}_2(\mathbb{Z}_3)$ has order 16. Which of the 14 listed groups of order 16 is it isomorphic to?

[1] These three matrices are known as the *Pauli Spin Matrices*

K Fibonacci Sequence, Pascal's Triangle, and more

For convenience I record some numbers which will come up over and over again.

K.1 Binomial Theorem

Theorem K.1 (Binomial Theorem). *Suppose that x and y are elements of some ring R which commute with each other. Then, for every positive integer n we have*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

First a lemma

Lemma K.2.

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$

Proof. The standard, easy proof of this involves a short direct calculation the formula. For a slightly different proof let's look at $\binom{n}{k}$ as the number of k element subsets of a set with n elements, which we take as $\overline{n}$. Well a $k+1$ element subset of $\overline{n+1}$ either contains $n+1$ or it doesn't. In the second case we could think of it as a $k+1$ element subset of $\overline{n}$, in the first we could think of it as a k element subset of $\overline{n}$ unioned with $\{n+1\}$. ∎

Proof. (Proof of binomial theorem) We proceed by induction. The base case of works since $\binom{0}{0} = 1$. Now assume the result is true for n. Then

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)^n \cdot (x+y) \\
&= \left(\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}\right)(x+y) \\
&= \left(\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}\right) \cdot x + \left(\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}\right) \cdot y && \text{(Distributive prop)} \\
&= \sum_{k=0}^{n} \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k} && (x \text{ and } y \text{ commute}) \\
&= \sum_{k=1}^{n+1} \binom{n}{k-1} x^k y^{n+1-k} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k} && \text{(Reindexing sum)} \\
&= \sum_{k=0}^{n+1} \left(\binom{n}{k-1} + \binom{n}{k}\right) x^k y^{n+1-k} && (\text{Check } k=0, k=n+1 \text{ cases}) \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k} && \text{(By the lemma)} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k}.
\end{aligned}$$

which is what we wanted. ∎

Let p be a prime. Suppose for a ring R we had $p \cdot 1_R = 0_R$. Then we say that the ring R has *characteristic* p. Note that this implies the for all $r \in R$, $p \cdot r = 0_R$ as well since

$$p \cdot r = p \cdot (1_R \cdot r) = (p \cdot 1_R) \cdot r = 0_R \cdot r = 0_R.$$

For example, clearly $\mathbb{Z}_p$ has characteristic p. But also so does the ring of matrices with coefficients in $\mathbb{Z}_p$ or any degree, and the ring of polynomials with coefficients in $\mathbb{Z}_p$.

Theorem K.3 (Freshmen's Dream). *Suppose that R is a ring with characteristic p. Then for any pair of commuting elements, x and y, we have*

$$(x+y)^p = x^p + y^p.$$

Proof. For starters, note that for any integer $1 < k < p$, we have that the integers $k!$ and $(p-k)!$ do not divide p. Therefore since $\binom{p}{k}$ is an integer, this proves that p divides it.

Now using the binomial theorem, we have that

$$(x+y)^p = y^p + \sum_{k=1}^{p-1} \binom{p}{k} x^k y^{p-k} + x^p.$$

But we just explained that if R has characteristic p, then p times anything in R is zero, but since p divides $\binom{p}{k}$ for all $1 < k < p$, we have that each term in the middle sum must be zero, and so the entire middle sum must also be zero. Hence we are left just with $x^p + y^p$ as desired. ∎

EXERCISE K.1. Let a, b, p, i and j be integers with p prime and $i, j > 0$. Prove that

$$(a + bp^i)^{p^j} \equiv_{p^{i+j}} a^p.$$

SOLUTION

Hint! Easy.

EXERCISE K.2.

SOLUTION

(1) Prove that

$$\sum_{m=0}^{n} \binom{n}{m} = 2^n.$$

Hint! Expand $(1 + x)^n$ using the binomial theorem and then plug in 1 for x.

(2) Prove that

$$\sum_{m=0}^{n} (-1)^m \binom{n}{m} = 0.$$

Hint! Expand $(1 + x)^n$ using the binomial theorem and then plug in -1 for x.

(3) Prove that

$$\sum_{m \text{ even}}^{n} (-1)^m \binom{n}{m} = 2^{n-1}.$$

Hint! Add the results from the first two parts.

(4) Prove that

$$\sum_{m \text{ odd}}^{n} (-1)^m \binom{n}{m} = 2^{n-1}.$$

Hint! Subtract the results from the first two parts.

EXERCISE K.3. Let p be a prime and m a positive integer.

SOLUTION

(1) Suppose an integer m written in base β is

$$(a_n.a_{n-1}.\cdots.a_1.a_0)_\beta.$$

Prove that

$$\eta_p(m) = \frac{m - \mathrm{S}_\beta(m)}{p-1}.$$

where $\mathrm{S}_\beta(m)$ is the sum of the digits, i.e.

$$\mathrm{S}_\beta(m) = \sum_{i=0}^{n} a_i$$

{Does this fit? Already given this exercise?}

(2) For positive inters integers a and b prove that

$$\eta_p\left(\left(\begin{matrix} p^{a+b} \\ p^b \end{matrix}\right)\right) = a.$$

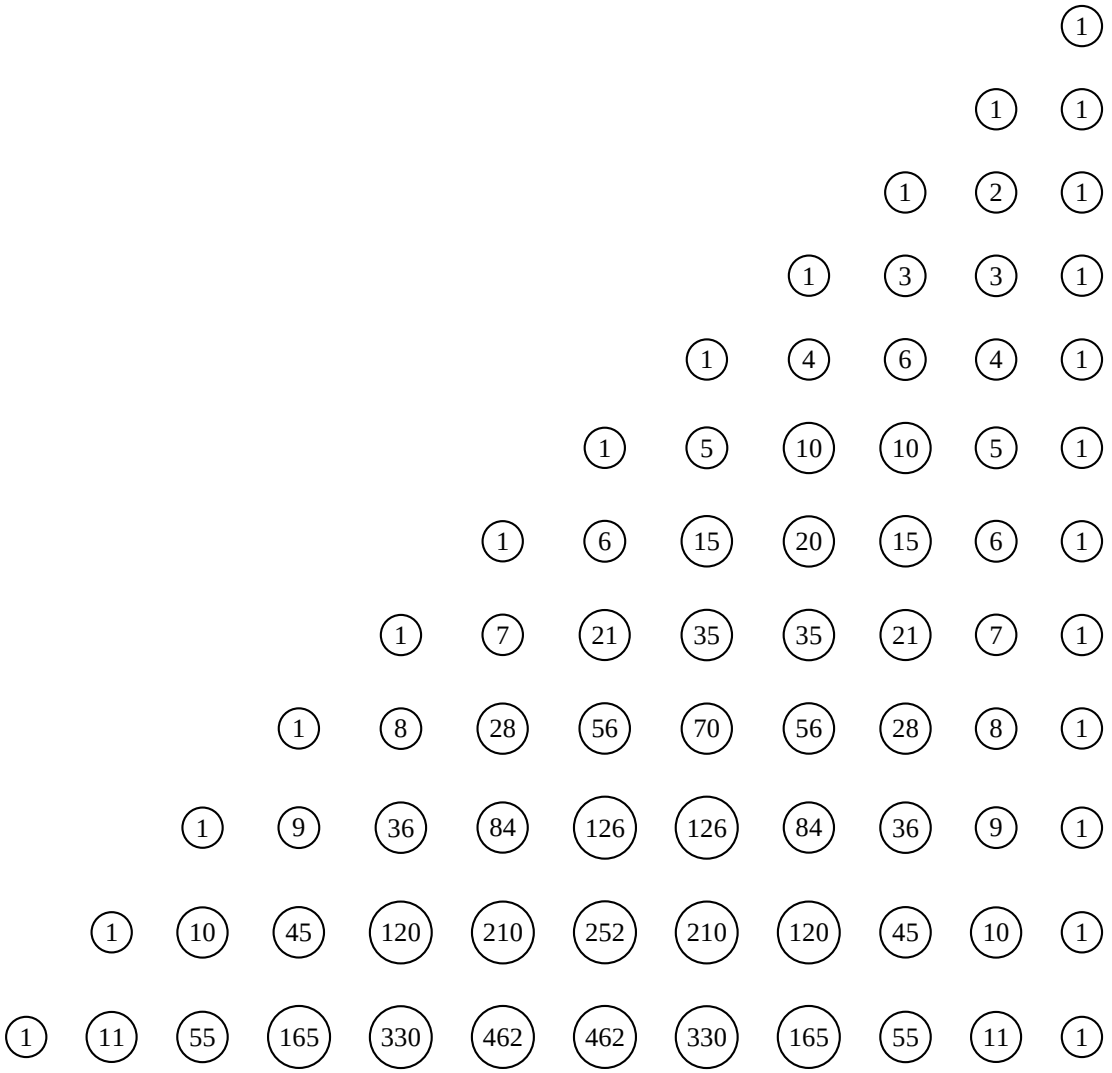
K.2 Partitions of n

For convience, I rerecord the table from page 365.

n	1	2	3	4	5	6	7	8	9	10
$\pi(n)$	1	2	3	5	7	11	15	22	30	42
n	11	12	13	14	15	16	17	18	19	20
$\pi(n)$	56	77	101	135	176	231	297	385	490	627
n	21	22	23	24	25	26	27	28	29	30
$\pi(n)$	792	1002	1255	1575	1958	2436	3010	2718	4565	5604
n	31	32	33	34	35	36	37	38	39	40
$\pi(n)$	6842	8349	10143	12310	14883	17977	21637	26015	31185	37338
n	41	42	43	44	45	46	47	48	49	50
$\pi(n)$	44583	53174	63261	75175	89134	105558	124754	147273	173525	204226

Table K.1: Partitions of n

K.3 Pascal's Triangle



K.4 Trace Polynomials

Define

$$\mathbf{G}_t \stackrel{\text{def}}{=} \begin{pmatrix} t & -1 \\ 1 & 0 \end{pmatrix}.$$

and

$$u_n(t) \stackrel{\text{def}}{=} \operatorname{tr}(\mathbf{G}_t^n).$$

As long as θ is not a integer multiple of π the standard rotation matrix

$$\varrho_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

whose n^{th} power is

$$\varrho_\theta^n = \varrho_{n\theta} = \begin{pmatrix} \cos n\theta & \sin n\theta \\ -\sin n\theta & \cos n\theta \end{pmatrix} \tag{1}$$

is conjugate[I] in $\mathbf{GL}_2(\mathbb{R})$ to the matrix

$$\mathbf{G}_{2\cos\theta} := \begin{pmatrix} 2\cos\theta & -1 \\ 1 & 0 \end{pmatrix}.$$

Therefore, just by looking at (1), we see the trace of ϱ_θ^n is $2\cdot\cos n\theta$, we have that

$$u_n(2\cos\theta) = \operatorname{tr}\left(\mathbf{G}_{2\cos\theta}^n\right) = 2\cdot\cos n\theta$$

as well.

Another application of this, any matrix with determinant 1 and trace t is conjugate to $\mathbf{G}_t$ (again, they have the same trace).

THEOREM K.4. *For any 2 by 2 matrix* $\mathbf{A}$ *we have*

$$u_n(\operatorname{tr}(\mathbf{A})) = \operatorname{tr}(\mathbf{A}^n).$$ ■

Therefore we can kill two birds with one stone and just computer the polynomials $u_n(x)$.

This is because a generic matrix of trace t and determinant 1 is conjugate to the matrix (again, same determinant and trace).

Using induction, one can easily prove by induction that

$$\mathbf{G}_t^n = \begin{pmatrix} a_n & a_{n-1} \\ -a_{n-1} & -a_{n-2} \end{pmatrix}$$

where a_n is the degree n polynomial defined recursively via

$$a_{-1} = 0 \qquad a_0 = 1 \quad \text{and} \quad a_n = t\cdot a_{n-1} - a_{n-2}.$$

Note, inductively assuming that a_{n-1} is monic of degree $n-1$ and a_{n-2} is monic of degree $n-2$ we see that $t\cdot a_{n-1}$ is monic of degree n, and since a_{n-2} has degree $n-2$, the difference $t\cdot a_{n-1} - a_{n-2}$ does not change the leading term of $t\cdot a_{n-1}$ proving that for each $n\in\mathbb{N}$, $a_n(t)$ is monic of degree n.

Also note that with this notation,

$$u_n(t) = \operatorname{tr}(\mathbf{G}_t^n) = a_n(t) - a_{n-2}(t) = t\cdot a_{n-1}(t) - 2\cdot a_{n-2}(t).$$

This proves

[I] They have the same determinant and trace and neither are scalar (as long as θ is not a multiple of π). Actually, to be more concrete one can verify that if we define

$$\mathbf{C} := \begin{pmatrix} \cos\theta & 1 \\ \sin\theta & 0 \end{pmatrix},$$

(which is invertible as long as $\sin\theta \neq 0$, i.e. as long as θ is not an integer multiple of π) then a direct check shows us that

$$\varrho_\theta = \mathbf{C}\cdot\mathbf{G}_{2\cos\theta}\cdot\mathbf{C}^{-1}.$$

PROPOSITION K.5. *For each $n \in \mathbb{N}$, the polynomials $a_n(t)$ and $u_n(t)$ are monic of degree n.* ■

COROLLARY K.6. *For each integer n,*

(1) *$2\cos n\theta$ is a monic polynomial in $2\cos\theta$;*

(2) *For any 2 by 2 matrix $\mathbf{A}$ of determinant 1, $\operatorname{tr}(\mathbf{A}^n)$ is a monic polynomial in $\operatorname{tr}(\mathbf{A})$.* ■

THEOREM K.7. *For any rational number, $\frac{m}{n}$, $2\cos\frac{m\pi}{n}$ is the root of some monic polynomial, i.e. it is an algebraic integer.*

PROOF. We have

$$u_n\left(2\cos\frac{m\pi}{n}\right) = 2\cos\frac{mn\pi}{n} = 2\cos m\pi = 2\cdot(-1)^m.$$

Therefore, if m is even, then $2\cos\frac{m\pi}{n}$ is a root of the polynomial

$$u_n(t) - 2$$

and if m is odd it is the root of the polynomial

$$u_n(t) + 2.$$

In either case, the polynomial in question is clearly monic of degree n. ■

We showed that it's a root of a degree n polynomial. Actually this is wildly inefficient. It is known that in general, for any reduced fraction $\frac{m}{n}$, $2\cos\frac{m\pi}{n}$ is a root of monic polynomial of degree $\vartheta(n)/2$. Actually, this is something one can prove easily after their first semester as a graduate student immediately from some basic Galois theory.

Let's do some calculations. Using our recursion formulas, we quickly see that

$$\mathbf{G}_t^1 = \begin{pmatrix} t & 1 \\ -1 & 0 \end{pmatrix} \qquad \mathbf{G}_t^2 = \begin{pmatrix} t^2-1 & t \\ -t & -1 \end{pmatrix} \qquad \mathbf{G}_t^3 = \begin{pmatrix} t^3-2t & t^2-1 \\ -(t^2-1) & -t \end{pmatrix}$$

$$\mathbf{G}_t^4 = \begin{pmatrix} t^4-3t^2+1 & t^3-2t \\ -(t^3-2t) & -(t^2-1) \end{pmatrix} \qquad \mathbf{G}_t^5 = \begin{pmatrix} t^5-4t^3+3t & t^4-3t^2+1 \\ -(t^4+3t^2+1) & -(t^3-2t) \end{pmatrix}$$

$$\mathbf{G}_t^6 = \begin{pmatrix} t^6-5t^4+6t^2-1 & t^5-4t^3+3t \\ -(t^5+4t^3+3t) & -(t^4-3t^2+1) \end{pmatrix} \qquad \mathbf{G}_t^7 = \begin{pmatrix} t^7-6t^5-10t^3+4t & t^6-5t^4+6t^2-1 \\ -(t^6-5t^4+6t^2-1) & -(t^5-4t^3+3t) \end{pmatrix}$$

$$\mathbf{G}_t^8 = \begin{pmatrix} t^8-7t^6+15t^4-10t^2+1 & t^7-6t^5-10t^3+4t \\ -(t^7-6t^5-10t^3+4t) & -(t^6-5t^4+6t^2-1) \end{pmatrix}$$

$$\mathbf{G}_t^9 = \begin{pmatrix} t^9-8t^7+21t^5-20t^3+5t & t^8-7t^6+15t^4-10t^2+1 \\ -(t^8-7t^6+15t^4-10t^2+1) & -(t^7+6t^5+10t^3+4t) \end{pmatrix}$$

$$\mathbf{G}_t^{10} = \begin{pmatrix} t^{10}-9t^8+28t^6-35t^4+15t^2-1 & t^9-8t^7+21t^5-20t^3+5t \\ -(t^9-8t^7+21t^5-20t^3+5t) & -(t^8-7t^6+15t^4-10t^2+1) \end{pmatrix},$$

etc. Taking traces of these matrices shows us that

$$u_0(t) = 2 \qquad u_1(t) = t$$
$$u_2(t) = t^2-2 \qquad u_3(t) = t^3-3t$$
$$u_4(t) = t^4-4t^2+2 \qquad u_5(t) = t^5-5t^3+5t$$
$$u_6(t) = t^6-6t^4+9t^2-2 \qquad u_7(t) = t^7-7t^5+14t^3-7t$$
$$u_8(t) = t^8-8t^6+20t^4-16t^2+2 \qquad u_9(t) = t^9-9t^7+27t^5-30t^3+9t$$
$$u_{10}(t) = t^{10}-10t^8+35t^6-50t^4+25t^2-2 \qquad u_{11}(t) = t^{11}-11t^9+44t^7-77t^5+55t^3-11t$$
$$u_{12}(t) = t^{12}-12t^{10}+54t^8-112t^6+105t^4-36t^2+2 \qquad u_{13}(t) = t^{13}-13t^{11}+65t^9-156t^7+182t^5-91t^3+13t.$$

Exercise K.4. Prove that $u_n \circ u_m = u_{n \cdot m}$.
Hint! This is obvious provided you are looking at things the correct way.

SOLUTION

Exercise K.5. Prove that $u_n(x)$ and $a_n(x)$ are even when n is even and are odd when n is odd. [II]

SOLUTION

Exercise K.6. Prove that $u_n(x)_{n-1} = 0$.
Hint! Use the previous exercise to save some time.

SOLUTION

Exercise K.7. Prove that $u_n(x)_{n-2}$ is $-n$.

SOLUTION

Note that for any polynomial $p(x)$, $p(0)$ is the constant coefficient.

Exercise K.8.

(1) $u_n(0) = \mathbf{Re}(2 \cdot i^n)$.

(2) $a_n(0) = \frac{1}{2} u_{n-1}(0)$.

SOLUTION

Exercise K.9. Prove that $(u_n)_1(x) = \mathbf{Re}\, i^{n-1} \cdot n$.

SOLUTION

Exercise K.10. Prove that $(u_n)_2(x) = \mathbf{Re}\, i^{n-2} \cdot n^2$.

SOLUTION

[II] If you are not familiar with the definition of an even or odd function, you should be able to figure them out just from looking at the what you get.

EXERCISE K.11. If p is prime, prove that all coefficients, with the exception of the leading coefficient, of $u_p(x)$ are divisible by p.
Hint! Follow an earlier proof to prove that

$$\operatorname{tr}\left(\mathbf{G}_t^p\right) \equiv_p \operatorname{tr}\left(\mathbf{G}_t\right)^p .$$

SOLUTION

Careful! The exact same proof doesn't quite work since we previously assumed that $t \in \mathbb{Z}_p$. This is no longer the case since t is a variable. However, a slight modification of the proof which said that $\operatorname{tr}(\mathbf{A}^p) = \operatorname{tr}(\mathbf{A})$ for any matrix in $\mathbf{SL}_2(\mathbb{Z}_2)$ will prove this more general exercise.

Note that for all integers k and n we have

$$u_n\left(2\cos\frac{k\pi}{n}\right) = 2\cos\frac{nk\pi}{n} = 2\cos k\pi = 2\cdot(-1)^k.$$

Using the trig identity $\cos(\pi-\theta) = -\cos\theta$, we see that all of these guys are roots of the polynomial

$$u_n(t) - 2.$$

The problem is that these numbers are not distinct. Indeed, the polynomial $u_n(t)-2$ is always reducible and there polynomials $v_n(t)$ so that

$$u_n(t) - 2 = \begin{cases} (t-2)v_n^2(t) \text{ if } n \text{ is odd} \\ (t-2)(t+2)v_n^2(t) \text{ if } n \text{ is odd.} \end{cases}$$

We will not prove this, but instead we back up a bit to take a second look at the polynomials $a_n(x)$. Note that since $\sin n\pi = 0$ for all integers, we know that the matrix $\varrho_{n\pi}$ is a scalar matrix (actually equal to $\pm\mathbf{I}_2$). Since the matrix $\mathbf{G}_{2\cos\frac{k\pi}{n}}$ is conjugate to $\varrho_{n\pi}$ it too must be this same scalar matrix. Therefore, its off diagonal term $a_{n-1}(2\cos\frac{k\pi}{n}) = 0$. This gives us a degree $n-1$-polynomial that we've found all of the roots to:

$$\left\{2\cos\frac{k\pi}{n} : 1 \le k < n\right\}.$$

(Note that none of this works for k a multiple of n, $k = qn$, since in this special case the matrices $\mathbf{G}_{2\cos q\pi}$ and $\varrho_{q\pi}$ are bf not conjugate to each other since the first is diagonal while the second is not.

This means the polynomial $a_{n-1}(x)$ factors as

$$a_{n-1}(x) = \prod_{i=1}^{n-1}\left(x - 2\cdot\cos\frac{k\pi}{n}\right).$$

Modulo some annoying details that I leave to you, we have proven that

PROPOSITION K.8. *For n odd*

$$\prod_{k=1}^{n-1}(-1)^k\left(2\cos\frac{k\pi}{n}\right) = 1. \quad \blacksquare$$

This means that for n odd, the algebraic integer $2\cos\frac{k\pi}{n}$ is invertible for every integer. If n is even and not a power of 2, then we can

K.5 Fibonacci Sequence

n	F_n	Prime Factorization	n	F_n	Prime Factorization
1	1	1	31	1,346,269	$557 \cdot 2417$
2	1	1	32	2,178,309	$3 \cdot 7 \cdot 47 \cdot 2207$
3	2	2	33	3,524,578	$2 \cdot 89 \cdot 19801$
4	3	3	34	5,702,887	$1597 \cdot 3571$
5	5	5	35	9,227,465	$5 \cdot 13 \cdot 141961$
6	8	2^3	36	14,930,352	$2^4 \cdot 3^3 \cdot 17 \cdot 19 \cdot 107$
7	13	13	37	24,157,817	$73 \cdot 149 \cdot 2221$
8	21	$3 \cdot 7$	38	39,088,169	$37 \cdot 113 \cdot 9349$
9	34	$2 \cdot 17$	39	63,245,986	$2 \cdot 233 \cdot 135721$
10	55	$5 \cdot 11$	40	102,334,155	$3 \cdot 5 \cdot 7 \cdot 11 \cdot 41 \cdot 2161$
11	89	89	41	165,580,141	$2789 \cdot 59369$
12	144	$2^4 \cdot 3^2$	42	267914296	$2^3 \cdot 13 \cdot 29 \cdot 211 \cdot 421$
13	233	233	43	433494437	433494437
14	377	$13 \cdot 29$	44	701408733	$3 \cdot 43 \cdot 89 \cdot 199 \cdot 307$
15	610	$2 \cdot 5 \cdot 61$	45	1134903170	$2 \cdot 5 \cdot 17 \cdot 61 \cdot 109441$
16	987	$3 \cdot 7 \cdot 47$	46	1836311903	$139 \cdot 461 \cdot 28657$
17	1,597	1597	47	2971215073	2971215073
18	2,584	$2^2 \cdot 17 \cdot 19$	48	4807526976	$2^6 \cdot 3^2 \cdot 7 \cdot 23 \cdot 47 \cdot 1103$
19	4,181	$37 \cdot 113$	49	7778742049	$13 \cdot 97 \cdot 6168709$
20	6,765	$3 \cdot 5 \cdot 11 \cdot 41$	50	12586269025	$5^2 \cdot 11 \cdot 101 \cdot 151 \cdot 3001$
21	10,946	$2 \cdot 13 \cdot 421$	51	20365011074	$2 \cdot 1597 \cdot 6376021$
22	17,711	$89 \cdot 199$	52	32951280099	$3 \cdot 233 \cdot 521 \cdot 90481$
23	28,657	28657	53	53316291173	$953 \cdot 55945741$
24	46,368	$2^4 \cdot 3^2 \cdot 7 \cdot 23$	54	86267571272	$2^3 \cdot 17 \cdot 19 \cdot 53 \cdot 109 \cdot 5779$
25	75,025	$5^2 \cdot 3001$	55	139583862445	$5 \cdot 89 \cdot 661 \cdot 474541$
26	121,393	$233 \cdot 521$	56	225851433717	$3 \cdot 7^2 \cdot 13 \cdot 29 \cdot 281 \cdot 14503$
27	196,418	$2 \cdot 17 \cdot 53 \cdot 109$	57	365435296162	$2 \cdot 37 \cdot 113 \cdot 797 \cdot 54833$
28	317,811	$3 \cdot 13 \cdot 29 \cdot 281$	58	591286729879	$59 \cdot 514229 \cdot 19489$
29	514,229	514229	59	956722026041	$353 \cdot 2710260697$
30	832,040	$2^3 \cdot 5 \cdot 11 \cdot 31 \cdot 61$	60	1548008755920	$2^4 \cdot 3^2 \cdot 5 \cdot 11 \cdot 31 \cdot 41 \cdot 61 \cdot 2521$

K.6 Möbius Inversion Formula

Let $\mathbb{C}^{\mathbb{N}}$ be short-hand for the set of all functions $\mathbb{N} \to \mathbb{C}$. We turn $\mathbb{C}^{\mathbb{N}}$ into an abelian monoid by defining a product of $f, g \in \mathbb{C}^{\mathbb{N}}$

$$(f \star g)(n) \stackrel{\text{def}}{=} \sum_{d|n} f\left(d\right) g\left(\frac{n}{d}\right).$$

We can rewrite this in a more symmetric way as follows:

$$(f \star g)(n) = \sum_{ab=n} f(a)g(b).$$

This makes it clear that $(\mathbb{C}^{\mathbb{N}}, \star)$ is abelian.

Note that both $(f \star g) \star h$ and $f \star (g \star h)$ equal

$$n \mapsto \sum_{abc=n} f(a)g(b)h(c).$$

Therefore

$$(f \star g) \star h = f \star (g \star h)$$

$(\mathbb{C}^{\mathbb{N}}, \star)$ is associative.

Lastly, if we define $e \in \mathbb{C}^{\mathbb{N}}$ as

$$e(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n \neq 1 \end{cases}$$

then since for any $f \in \mathbb{C}^{\mathbb{N}}$, we have

$$(f \star e)(n) = \sum_{ab=n} f(a)e(b) = \sum_{ab=n, b=1} f(a) = \sum_{a=n} f(a) = f(n)$$

we see that e is the identity of $(\mathbb{C}^{\mathbb{N}}, \star)$. Therefore we have proved

Proposition K.9. *$(\mathbb{C}^{\mathbb{N}}, \star)$ is a monoid.* ■

Definition K.10. The function $\mathrm{I}\colon \mathbb{N} \to \mathbb{C}$ is defined to be the constant function $\mathrm{I}(n) = 1$ for all $n \in \mathbb{N}$.

Notice that for any $f\colon \mathbb{N} \to \mathbb{C}$ we have that

$$(f \star \mathrm{I})(n) = \sum_{d|n} f(d)\, \mathrm{I}\left(\frac{n}{d}\right) = \sum_{d|n} f(d).$$

Definition K.11. The *Möbius function*, $\mu\colon \mathbb{N} \to \mathbb{C}$ is defined inductively as follows:

(1) $\mu(1) = 1$

(2) Suppose $\mu(n)$ is defined and let p be a prime. Then

$$\mu(np) := \begin{cases} (-1)\mu(n) & \text{if } p \nmid n \\ 0 & \text{if } p \mid n. \end{cases}$$

Therefore, if

$$N = p_1^{n_1} \cdots p_m^{n_m}$$

then

$$\mu(N) := \begin{cases} 0 & \text{if } n_i > 1 \text{ for some prime } i \in \overline{m} \\ (-1)^m & \text{otherwise.} \end{cases}$$

This means that $\mu(N) \neq 0$ iff N is square free. The first values of μ are:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\mu(n)$	1	−1	−1	0	−1	1	−1	0	0	1	−1	0	−1	1	1

.

LEMMA K.12. I *is an inverse to* μ *in* $(\mathbb{C}^{\mathbb{N}}, \star)$.

PROOF. We have that

$$(\mu \circ \mathrm{I})(n) = \sum_{d|n} \mu(d).$$

Write n as

$$n = p_1^{n_1} \cdots p_m^{n_m}.$$

Let P denote the set of prime divisors of n,

$$P = \{p_1, \ldots, p_m\}.$$

For a subset of P, $\overline{P}$, define $d_{\overline{P}}$ as the product of the elements (primes) $\overline{P}$. We also define $d_\varnothing$ to be 1. Note that each $d_{\overline{P}} \mid n$ and every divisor d that does not have this form is divisible by the same prime more than once and so we have that $\mu(d) = 0$ for such divisors. Of course

$$\mu(d_{\overline{P}}) = (-1)^{\#\overline{P}}. \tag{2}$$

Since $\mu(1) = 1$, we know that $\sum_{d|1} \mu(d) = 1$. Now for $n \neq 1$ we see that

$$\begin{aligned}
\mathrm{I} \star \mu &= \sum_{d|n} \mu(d) \\
&= \sum_{\substack{d|n \\ \text{square free}}} \mu(d) && (\mu(d) = 0 \text{ if } d \text{ is not square free}) \\
&= \sum_{\overline{P} \subseteq P} \mu(d_{\overline{P}}) && (\text{All sf divisors of } n \text{ have form } d_{\overline{P}} \text{ for some } \overline{P} \subseteq P.) \\
&= \sum_{k=0}^{m} \sum_{\overline{P} \in \mathcal{P}_k(P)} \mu(d_{\overline{P}}) && (\mathcal{P}_k(P) = k \text{ element subsets of } P) \\
&= \sum_{k=0}^{m} \sum_{\overline{P} \in \mathcal{P}_k(P)} (-1)^k && (\text{Equation (2)}) \\
&= \sum_{k=0}^{m} (-1)^k \binom{n}{k} && (\#\mathcal{P}_k(P) = \binom{n}{k}) \\
&= 0. && (\text{See Exercise K.2(2)})
\end{aligned}$$

Therefore

$$\mathrm{I} \star \mu = \mu \star \mathrm{I} = e$$

as desired. ■

THEOREM K.13 (Möbius Inversion Theorem)**.** *Suppose that*

$$\widehat{f}(n) = \sum_{d|n} f(d)$$

for some $f \in \mathbb{C}^{\mathbb{N}}$. *Then*

$$f(n) = \sum_{d|n} \mu(d) \widehat{f}(n/d).$$

PROOF. The hypothesis that

$$\widehat{f}(n) = \sum_{d|n} f(d)$$

is another way of saying that

$$\widehat{f} = f \star \mathrm{I}.$$

But the above lemma says that μ is an inverse to I (under the operation $\star$). Therefore

$$\begin{aligned}
\widehat{f} \star \mu &= (f \star I) \star \mu \\
&= f \star (I \star \mu) \\
&= f \star e \\
&= f.
\end{aligned}$$

Plugging an integer n into this equality is the statement

$$f(n) = \sum_{d|n} \mu(d) \widehat{f}(n/d).$$

as desired. ∎

EXAMPLE K.14. In any group G we have that

$$\mathbf{O}_n(G) = \sum_{d|n} \mathbf{o}_d(G).$$

Hence

$$\mathbf{o}_n(G) = \sum_{d|n} \mu\left(\frac{n}{d}\right) \mathbf{O}_d\left(G\right).$$

EXAMPLE K.15. In the specific case that G is the cyclic group $\mathbb{Z}_n$, we have that for every $m|n$, $\mathbf{o}_m(G) = \vartheta(m)$ and $\mathbf{O}_m(G) = m$. Therefore the formula reads

$$\vartheta(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) d.$$

EXAMPLE K.16. Define $\tau(n)$ to be the number of divisors of the number n, that is

$$\tau(n) = \#\{d : d \mid n\}.$$

Then

$$\tau(n) = \sum_{d|n} 1 = \sum_{d|n} \mathrm{I}(d).$$

Hence

$$1 = \sum_{d|n} \tau\left(d\right) \mu\left(\frac{n}{d}\right).$$

L SOLUTIONS

2.1 (2.1) Assume $\mathcal{A}$ is a set.

(1) Assume $\mathcal{A} \in \mathcal{A}$. Then, since $\mathcal{A} = \{A : A \text{ is a set and } A \notin A\}$ we must have that

$$\mathcal{A} \text{ is a set and } \mathcal{A} \notin \mathcal{A}.$$

Therefore $\mathcal{A} \notin \mathcal{A}$.

(2) Assume on the other hand that $\mathcal{A} \notin \mathcal{A}$. Then, by the first assumption $\mathcal{A}$ is a set and by this latest assumption $\mathcal{A} \notin \mathcal{A}$. These are the two things needed for $\mathcal{A} \in \mathcal{A}$.

(3) For any set X and any object x, the statement $x \in X$ is either true or false, and since we have just proven that the statement $\mathcal{A} \in \mathcal{A}$ is neither true nor false, $\mathcal{A}$ cannot be a set.

2.2 (2.2) We know that for any set X that $\varnothing_P \subseteq X$ and $\varnothing_{P'} \subseteq X$. Thus taking X as $\varnothing_{P'}$ in the first statement shows that $\varnothing_P \subseteq \varnothing_{P'}$. On the other hand using this in the second statement shows us that $\varnothing_{P'} \subseteq \varnothing_P$ as well. Since we have mutual inclusion, $\varnothing_P = \varnothing_{P'}$.

2.3 (2.3)

2.4 (2.4) Like any set, we get one inclusion $\varnothing \subseteq \varnothing \times A$ for free. We need to prove that $\varnothing \times A \subseteq \varnothing$. Indeed, if it were false there would be an object in $\varnothing \times A$ which is not in $\varnothing$, i.e. there would be a point in $\varnothing \times A$. But this is impossible since

$$(x, y) \in \varnothing \times A \Rightarrow x \in \varnothing \wedge y \in A \Rightarrow x \in \varnothing$$

which is false since the empty set contains no objects.

∅ IS CONTAINED IN ALL SETS!

2.5 (2.5)

2.6 (2.6) No notion B/B is meaningless and certainly cannot be used to prove anything.

2.7 (2.7) Typically (a, b) is defined as the set $\{a, \{a, b\}\}$. For instace, see [SSA01].

2.8 (2.8) You can obviously do this by induction. However, this is proven most easily by noting that in the sum

$$1 + 2 + \cdots (n-2) + (n-1) \tag{1}$$

we can rearrange terms so that we have

$$[1 + (n-1)] + [2 + (n-2)] + [3 + (n-3)] + \cdots = n + n + n + \cdots . \tag{2}$$

- If n is odd (so that $n-1$ is even) there are exactly half as many terms in the sum on the RHS of (2) as there were in (1), i.e. there are $(n-1)/2$ terms in the sum. As such

$$\begin{aligned}
\sum_{i=1}^{n-1} i = &= \underbrace{n + n + n + \cdots}_{\frac{n-1}{2}\text{-times}} && \text{(Using (2))} \\
&= \frac{(n-1)}{2} \cdot n \\
&= \frac{n \cdot (n-1)}{2} \\
&= \frac{n \cdot (n-1) \cdot (n-2)!}{2 \cdot (n-2)!} && \text{(Mult T\&B by } (n-2)!) \\
&= \frac{n!}{2! \cdot (n-2)!} && \text{(Use } n! = n \cdot (n-1) \cdot (n-2)!,\ 2! = 2) \\
&= \binom{n}{2}. && \text{((2) on page 12)}
\end{aligned} \tag{3}$$

- If n is even (so that $n-1$ is odd) then the sum on the RHS of (2) leaves out $n/2$ and is

$$\begin{aligned}\underbrace{n+n+n+\cdots}_{\frac{n-2}{2}\text{-times}}+\frac{n}{2} &= \frac{n-2}{2}\cdot n+\frac{n}{2}\\ &= \frac{(n^2-2n)+n}{2}\\ &= \frac{n^2-n}{2}\\ &= \frac{n\cdot(n-1)}{2}.\\ &= \binom{n}{2}. \qquad \text{(Same argument as end of (3))}\end{aligned}$$

Therefore, we have verified the formula in every case.

2.11 (2.11) This is always false. For the first set, $\mathcal{P}_2(\mathbb{N})$ contains pairs of elements while $\mathcal{P}_3(\mathbb{N})$ contains triples of elements. Since something cannot be a triple and a pair we actually have that

$$\mathcal{P}_2(\mathbb{N}) \cap \mathcal{P}_3(\mathbb{N}) = \varnothing.$$

2.14 (2.14)

2.15 (2.15)

2.16 (2.16)

(1) **(a)** For two people p_1 and p_3, we have $p_1 \mathcal{R}\circ\mathcal{R}\, p_3$ iff there is a person p_2 so that $p_2 \mathcal{R}\, p_3$ and $p_1 \mathcal{R}\, p_2$, i.e. p_3 is a parent of p_2 and p_2 is a parent of p_1, in other words p_3 is a grandparent of p_1.

(b) $p_2 \mathcal{R}^{-1}\, p_1$ means that p_2 is a child of p_1.

(c) People are related under this relation iff they are the same person or if they have a child together.

(d) People are related under this relation iff they have a common parent; i.e. iff they are either the same person, are siblings or are step siblings.

(2) The relation $\mathcal{R}$ is defined everywhere since everybody has a parent.

(3) The relation is not surjective since not every person is a parent.

2.17 (2.17)

We only prove $\mathcal{R}\circ \mathbf{1}_X = \mathcal{R}$. The other equality is proven in essentially the same manner.

$\subseteq$ [I] Let $(x,z) \in \mathcal{R}\circ\mathbf{1}_X$ be arbitrary. Then we have

$$\begin{aligned}(x,y)\in \mathcal{R}\circ\mathbf{1}_X &\Rightarrow x\, \mathcal{R}\circ\mathbf{1}_X\, y && \text{(Notation change)}\\ &\Rightarrow \exists x'\in X \text{ s.t. } x\, \mathbf{1}_X\, x' \wedge x'\, \mathcal{R}\, y && \text{(Defn of composition)}\\ &\Rightarrow \exists x'\in X \text{ s.t. } x = x' \wedge x'\, \mathcal{R}\, y && \text{(Meaning of Id)}\\ &\Rightarrow \exists x'\in X \text{ s.t. } (x,y)=(x',y) \wedge x'\, \mathcal{R}\, y && (x=x' \text{ and } y=y)\\ &\Rightarrow \exists x'\in X \text{ s.t. } (x,y)=(x',y) \wedge (x',y)\in \mathcal{R} && \text{(Notation change)}\\ &\Rightarrow (x,y)\in\mathcal{R}.\end{aligned}$$

Since (x,y) was arbitrary we have proven that $\mathcal{R}\circ\mathbf{1}_X \subseteq \mathcal{R}$.

[I] We are a little more pedantic than necessary here. (This is good practice for these sorts of exercises!)

$\supseteq$ Let $(x, y) \in \mathcal{R}$ be arbitrary. We also have $(x, x) \in \mathbf{1}_X$. Using these two facts:

$$(x, y) \in \mathcal{R} \wedge (x, x) \in \mathbf{1}_X \Rightarrow (x, y) \in \mathcal{R} \circ \mathbf{1}_X.$$

Since (x, y) was arbitrary we have proven that $\mathcal{R} \circ \mathbf{1}_X \supseteq \mathcal{R}$.

2.18 (2.18)

2.19 (2.19)

2.20 (2.20)

2.21 (2.21)

2.22 (2.22)

2.29 (2.29)

2.30 (2.30) Clearly the statement is true if either X or Y has an infinite number of elements. Therefore we assume X and Y have n and m elements respectively. Moreover, by choosing dictionaries, we can assume that $X = \overline{n}$ and $Y = \overline{m}$. Doing this, for any function $f : X \to Y$ we have a unique n-tuple of integers between 1 and m:

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

and conversely for any such n-tuple determines a unique function $X \to Y$ by sending the integer i the i^{th} entry in the tuple. Therefore the set of functions $X \to Y$ are in bijection with

$$\underbrace{\overline{m} \times \cdots \times \overline{m}}_{n\text{-times}}$$ [II]

which has

$$\#\left(\underbrace{\overline{m} \times \cdots \times \overline{m}}_{n\text{-times}}\right) = \underbrace{\#\overline{m} \cdot \cdots \cdot \#\overline{m}}_{n\text{-times}} = \underbrace{m \cdot \cdots \cdot m}_{n\text{-times}} = m^n$$

elements.

2.32 (2.32)

$\Rightarrow$ Suppose $\mathrm{F} : X \to X$ is a projection onto A.

(a) By definition $\mathrm{F}(A) = A$, therefore it's certainly the case that $\mathrm{F}(A) \subseteq A$.

(b) Let $a \in A$. Then since the image of F is A, there exists an $x \in X$ so that $\mathrm{F}(x) = a$. But using the equation $\mathrm{F} \circ \mathrm{F} = \mathrm{F}$ to x shows us that

$$\mathrm{F}(a) = \mathrm{F}(\mathrm{F}(x)) = \mathrm{F}(x) = a.$$

$\Leftarrow$ Suppose that F satisfies both parts of the hypothesis. By part **(2)** for any $a \in A$, $\mathrm{F}(a) = a$. In particular, every $a \in A$ is in the image of F. This proves that $A \subseteq \mathrm{F}(X)$. But part **(1)** says that the opposite inclusion holds and so $A = \mathrm{F}(A)$.

We still need to prove that $\mathrm{F} \circ \mathrm{F} = \mathrm{F}$. We know that for any $x \in X$, $\mathrm{F}(x) \in A$ so there exists an $a \in A$ so that $\mathrm{F}(x) = a$. But part **(2)** says that $\mathrm{F}(a) = a$. Therefore

$$\mathrm{F}(\mathrm{F}(x)) = \mathrm{F}(a) = a = \mathrm{F}(x).$$

Since x was arbitrary this proves that $\mathrm{F} \circ \mathrm{F} = \mathrm{F}$ as desired.

[II] Often, you'll see $\mathcal{F}\mathrm{un}(X, Y)$ denote as Y^X since functions $X \to Y$ can be thought of as ordered "X–tuples" of elements of Y.

2.33 (2.33)

2.35 (2.35)

⇒ Clearly the identity has crossing number 0.

⇐ Conversely, suppose that σ has crossing number 0. Assume it is not the identity. Then, the set

$$\operatorname{Supp}\sigma \stackrel{\text{def}}{=} \{i \in \overline{n} : \sigma(i) \neq i\}$$

is non-empty, and so it contains a least letter least letter, α, so that $\sigma(\alpha) \neq \alpha$. This means that if j is a letter with $j < i$, then $\sigma(j) = j$. But since σ is a bijection there exists a letter β which maps to α; $\sigma(\beta) = \alpha$. Since σ is single-valued, and $\sigma(\alpha) \neq \alpha$ we must have that $\alpha \neq \beta$. Therefore $\beta \in \operatorname{Supp}\sigma$, and since $\beta \neq \alpha$, we have $\beta > \alpha$. Moreover, since $\sigma(\alpha) \neq \alpha$ and since σ is one-to-one, we also have that $\sigma(\sigma(\alpha)) \neq \sigma(\alpha)$, and so $\sigma(\alpha) \in \operatorname{Supp}\sigma$ and is distinct from α, therefore $\sigma(\alpha) > \alpha$. Therefore, we have proved that

$$\alpha < \beta \quad \text{and} \quad \sigma(\alpha) > \alpha(\beta) = \alpha.$$

Therefore α and β cross and, in conclusion, $\operatorname{Cross}\sigma \neq 0$.

2.36 (2.36) Denote the image of a permutation σ as follows $\sigma(i) = i'$. Suppose that (i, j) contributes to the crossing number of a permutation σ meaning that $i < j$ and $i' > j'$. But since $i = \sigma^{-1}(i')$ and $j = \sigma^{-1}(j)$, this means that the pair (j', i') contribute to the crossing number of σ^{-1}. This gives us a bijection from the set of pairs which add to the crossing number of σ to the set of pairs which add to the crossing number of σ^{-1}. Therefore we have proven that

$$\operatorname{Cross}(\sigma) = \operatorname{Cross}(\sigma^{-1})$$

which was our goal.

2.37 (2.37)

(1) The identity

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

has crossing number 0.

(2) The permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}$$

has crossing number 1 since only 1 and 2 cross.

(3) The permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}$$

has crossing number 1 since only 1 and 2 cross.

(4) The permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$$

has crossing number 2 since both 1 and 2 cross 3 (and there are no other crossings). Alternatively, the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

also has crossing number 2.

(5) The permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

has crossing number 3 since 1 crosses both 1 and 2 and 2 crosses 3 and there are no other crossings.

(6) The permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 4 & 3 \end{pmatrix}$$

has crossing number four since 1 crosses 2, 3, and 4; and 3 crosses 4 and there are no other crossings.

(7) The permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 4 & 2 \end{pmatrix}$$

has crossing number five since 1 crosses 2, 3, and 4; and 2 and 3 cross 4, and there are no other crossings.

(8) The permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

(which is the half-twist from Exercise K.11crossmax just below) has crossing number 6 since everybody crosses everybody.

2.38 (**2.38**) Consider the involution which takes

$$1 \mapsto n \qquad 2 \mapsto (n-1) \qquad 3 \mapsto (n-3) \qquad \cdots \qquad n-1 \mapsto 1 \qquad n \mapsto 1.$$

We denote this permutation by σ_m. (In the literature σ_m is often called the *half twist*.) Clearly we have that

$$i < j \Rightarrow n - i > n - j.$$

This means the letter 1 crosses in front of $n-1$ other letters, the letter 2 crosses in front of $n-2$ letters, etc. This means that the cross number of σ_m is

$$(n-1) + (n-2) + \cdots + 2 + 1 = 1 + 2 + \cdots (n-2) + (n-1) = \sum_{i=1}^{n-1} i = \binom{n}{2}.$$

Conversely this is clearly the largest the crossing number of a permutation can be since there are no other crossings which are possible.

2.40 (**2.40**)

2.46 (**2.46**) Yes. For one, we used the fact that f was defined everywhere to guarantee that the relation $\sim_f$ was reflexive. Moreover we used the fact that f was single valued to be able to prove transitivity of the relation $\sim_f$.

2.47 (**2.47**) Suppose that a is even. Then there exists an integer k so that $a = 2k$. Suppose further that $a \equiv_2 b$. Then, by definition, $2 \mid (a-b)$, i.e. there exists an integer q so that $2q = a - b$. Using these two facts together:

$$2q = a - b \Rightarrow 2q = 2k - b \Rightarrow b = 2(k-q) \Rightarrow b \text{ is even.}$$

Since a number is either even or odd (and cannot be both) this completes the proof of exercise.

2.61 (**2.61**)

2.63 (2.63)

2.64 (2.64)

(1) Clearly π_m^n is induced from the usual quotient function function

$$\begin{aligned}\pi_m &: \mathbb{Z} \to \mathbb{Z}_m \\ &: a \mapsto [\, a \,]_m.\end{aligned}$$

which takes an integer to its class mod m. This is simply because we have

$$\pi_m^n([\, a \,]_n) = \pi_m(a).$$

(2) This part of the problem is an iff so we break it up accordingly.

$\Rightarrow$ Suppose that $m \mid n$. To show that π_m^n is well defined we need to show that

$$a \equiv_n b \Rightarrow \pi_m(a) = \pi_m(b).$$

But since $\pi_m(a) = [\, a \,]_m$, and since $\pi_m(a) = \pi_m(b)$ is true iff $a \equiv_m b$, we really need to prove that

$$a \equiv_n b \Rightarrow a \equiv_m b.$$

Now that we have unwinded the problem, it is now easy. Let's start over so the logic is clear.

$$\begin{aligned} m \mid n \quad \wedge \quad a \equiv_n b &\Rightarrow m \mid n \quad \wedge \quad n \mid (a - b) && \text{(Def'n of } a \equiv_n b\text{)} \\ &\Rightarrow m \mid (a - b) && \text{(} \mid \text{ is } \textbf{Transitive}\text{)} \\ &\Rightarrow a \equiv_m b. && \text{(Def'n of } a \equiv_m b\text{)} \end{aligned}$$

$\Leftarrow$ Conversely, suppose that π_m^n is well defined. Then since

$$[\, 0 \,]_n = [\, n \,]_n$$

we must have that

$$\pi_m^n([\, 0 \,]_n) = \pi_m^n([\, n \,]_n).$$

But by definition, the left hand side of this equation is $[\, 0 \,]_m$ while the right hand side is $[\, n \,]_m$. Therefore, we have

$$[\, 0 \,]_m = [\, n \,]_m\,,$$

i.e. $0 \equiv_m n$, or $n \mid m$, as desired.

2.69 (2.69)

2.70 (2.70)

3.3 (3.3)

(1) Suppose d_1 and d_2 satisfy **gcd1** and **gcd2**. Then, since d_2 satisfies **gcd1** we know that $d_2 \mid a$ and $d_2 \mid b$. But since d_2 satisfies **gcd2** we know that this implies that $d_1 \mid d_2$. Similarly we know that $d_2 \mid d_1$. Therefore $d_1 = \pm d_2$ (via the hint).

(2) The proof of this logically the same as the proof of the first part.

3.4 (3.4) Suppose $d \mid p$ (with p and d positive integers). Since d is an integer and p is prime either $p \mid d$ or p and d are relatively prime. If $p \mid d$, then since $d \mid p$ by assumption this means that $d = p$. On the other hand, if p and d are relatively prime, then since $d \mid d$ (any number divides itself) and $d \mid n$ (by assumption) we must have that $d = 1$ (by definition of what it means for two numbers to be relatively prime).

3.8 (3.8)

3.9 (3.9)

3.10 (3.10)

3.13 (3.13) Writing the complex numbers $z_1 + z_2 i$ and $w_1 + w_2 i$ in terms of the original definition (as ordered pairs) our assumption translates to

$$(z_1, z_2) = (w_1, w_2).$$

But by definition of ordered pairs, Definition 2.8, we know that this means that $z_1 = w_1$ and $z_2 = w_2$ as desired.

3.14 (3.14) Not possible to give a solution.

3.15 (3.15) Since for every u we have

$$\cos 3u = 4\cos^3 u - 3\cos u.$$

Things turn out a bit neater if instead we consider

$$2\cos 3u = 8\cos^3 u - 6\cos u \Rightarrow f(3u) = f^3(u) - 3f(u)$$

where $f(u) = 2\cos u$.

We have, in particular, taking $u = 3x$

$$f(9x) = f^3(3x) - 3f(3x).$$

Now plugging back in our original equation, but this time taking $u = x$ shows us

$$f(9x) = f^3(3x) - 3f(3x) = \left(f^3(x) - 3f(x)\right)^3 - 3\left(f^3(x) - 3f(x)\right).$$

Expanding $\left(f^3(x) - 3f(x)\right)^3$ gives

$$f^9(x) - 3f^6(x)\cdot(3f(x)) + 3f^3(x)\cdot(9f^2(x) - 27f^3(x) = f^9(x) - 9f^7(x) + 27f^5(x) - 27f^3(x).$$

Not forgetting to subtract of the

$$3\left(f^3(x) - 3f(x)\right)$$

term chilling at the end shows us that

$$f(9x) = f^9(x) - 9f^7(x) + 27f^5(x) - 27f^3(x) - 3f^3(x) + 9f(x) = f^9(x) - 9f^7(x) + 27f^5(x) - 30f^3(x) + 9f(x).$$

Plugging back in $f(x) = 2\cos x$ gives us

$$2\cos(9x) = 2^9\cdot\cos^9(x) - 9\cdot 2^7\cos^7(x) + 27\cdot 2^5\cos^5(x) - 30\cdot 2^3\cos^3(x) + 18\cos(x).$$

We can cancel the 2, but besides that there's not any real evident simpilification.

$$\cos(9x) = \left(256\cos^8(x) - 576\cos^6(x) + 432\cos^4(x) - 120\cos^2(x) + 9\right)\cos(x).$$

3.19 (3.19) There are two ways of doing this. I do both (providing more detail in the first method).

Method 1 In my opinion, the faster way (and also the method that has a more direct generalization) is to use Theorem 3.58. Indeed, we have to prove

$$(z \cdot w)^n = 1 \quad \text{and} \quad \left(z^{-1}\right)^n$$

whenever z^n and w^n each equal one. But this is easy:

$$(z \cdot w)^n = z^n \cdot w^n = 1 \cdot 1 = 1,$$

and

$$\left(z^{-1}\right)^n = (z^n)^{-1} = 1^{-1} = \frac{1}{1} = 1.$$

Method 2 The other method to prove this is to write

$$z = \cos\frac{2a\pi}{n} + \sin\frac{2a\pi}{n}\, i \text{ and } w = \cos\frac{2b\pi}{n} + \sin\frac{2b\pi}{n}\, i$$

for some integers a and b. Then note that

$$z \cdot w = \cos\frac{2(a+b)\pi}{n} + \sin\frac{2(a+b)\pi}{n}\, i \text{ and } z^{-1} = \cos\frac{-2a\pi}{n} + \sin\frac{-2a\pi}{n}\, i.$$

Hence both are of the form

$$\cos\frac{2m\pi}{n} + \sin\frac{2m\pi}{n}\, i$$

for some integer m and are therefore vertices in $\{\{n\}\}$.

3.21 (3.21)

3.23 (3.23)

3.27 (3.27)

3.28 (3.28)

3.35 (3.35)

(1) 5:

(2) $\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\,\mathbf{i}$

(3) $\frac{1}{2} - \frac{1}{2}\,\mathbf{i} + \frac{1}{2}\,\mathbf{j} + \frac{1}{2}\,\mathbf{k}$

(4) $\frac{3}{7}\,\mathbf{i} + \frac{2}{7}\,\mathbf{j} - \frac{6}{7}\,\mathbf{k}$

(5) $\mathbf{i}$

(6) $\frac{\sqrt{2}}{2} - \frac{\sqrt{2}\mathbf{j}}{2}\,\mathbf{j}$

(7) $\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\,\mathbf{j}$

(8) $\frac{\phi}{2} + \frac{\phi^{-1}}{2}\,\mathbf{i} + \frac{1}{2}\,\mathbf{j}$

(9) -1

(10) $\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}\,\mathbf{i} + \frac{\sqrt{2}}{2}\,\mathbf{j} - \frac{\sqrt{2}}{2}\,\mathbf{k}$

The only part that requires anything short of a straight forward calculation is **(8)**:

$$\begin{aligned}
\left\| \frac{\phi}{2} + \frac{\phi^{-1}}{2}\mathbf{i} + \frac{1}{2}\mathbf{j} \right\|^2 &= \left(\frac{\phi}{2}\right)^2 + \left(\frac{\phi^{-1}}{2}\right)^2 + \left(\frac{1}{2}\right)^2 \\
&= \left(\frac{\phi}{2}\right)^2 + \left(\frac{\phi - 1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 \qquad (\phi^{-1} = \phi - 1) \\
&= \frac{1}{4}\left(\phi^2 + (\phi - 1)^2 + 1^2\right) \\
&= \frac{1}{4}\left((\phi + 1) + (\phi^2 - 2\phi + 1) + 1\right) \\
&= \frac{1}{4}\left((\phi + 1) + (\phi + 1 - 2\phi + 1) + 1\right) \\
&= \frac{1}{4}\left((\cancel{\phi} + 1) + (2 - \cancel{\phi}) + 1\right) \\
&= \frac{1}{4}(4) \\
&= 1.
\end{aligned}$$

Here's a table which collects everything:

q	$\|\|q\|\|$	$\mathbb{R}$	$\mathbb{R}^3$	$\mathbb{S}^3$	$\mathbb{S}^2$	$\mathbb{S}^1$
1.	5	*yes*	*no*	*no*	*no*	*no*
2.	1	*no*	*no*	*yes*	*no*	*yes*
3.	1	*no*	*no*	*yes*	*no*	*no*
4.	1	*no*	*yes*	*yes*	*yes*	*no*
5.	1	*no*	*yes*	*yes*	*yes*	*yes*
6.	1	*no*	*no*	*yes*	*no*	*no*
7.	1	*no*	*no*	*yes*	*no*	*no*
8.	1	*no*	*no*	*yes*	*no*	*no*
9.	1	*yes*	*no*	*yes*	*no*	*yes*
10.	21	*no*	*no*	*no*	*no*	*no*

3.36 (??)

$\Rightarrow$ Suppose that $\mathsf{q} \in \mathbb{S}^2$. Then

$$\begin{aligned}
\mathsf{q} \in \mathbb{S}^3 \quad &\text{and} \quad \mathsf{q} \in \mathbb{R}^3 \quad (\text{Defn of } \mathbb{S}^2) \\
\Rightarrow \mathsf{q}^{-1} = \overline{\mathsf{q}} \quad &\text{and} \quad \overline{\mathsf{q}} = -\mathsf{q}. \quad (\text{Ext/Thm } 3.92 \ + \ \text{Defn of } \mathbb{R}^3)
\end{aligned}$$

Putting these together gives us the forward direction.

$\Leftarrow$ Suppose that $\mathsf{q}^{-1} = -\mathsf{q}$. Then we have

$$\mathsf{q}^2 = -(\mathsf{q} * (-\mathsf{q})) = -1.$$

But since in general $\mathsf{q} * \mathsf{q}^{-1} = \|\mathsf{q}\|^2$ we know that $\|\mathsf{q}\| = 1$. Therefore $\mathsf{q} \in \mathbb{S}^3$. It remains to show that q is pure. But we know that for $\mathsf{q} \in \mathbb{S}^3$, $\overline{\mathsf{q}} = \mathsf{q}^{-1}$ so combining this with the assumption we see that $\overline{\mathsf{q}} = -\mathsf{q}$, i.e. that q is pure.

3.45 (3.45)

(1) Write q as $q_R + \mathbf{q}_P$ Using the formula, we know that

$$\mathsf{q}^2 = (q_R{}^2 - \overbrace{\mathbf{q}_P \bullet \mathbf{q}_P}^{\|\mathbf{q}_P\|^2}) + (2q_R\mathbf{q}_P). \tag{4}$$

But since $\mathsf{q} \in \mathbb{S}^3$ we know that

$$\|\mathsf{q}\|^2 = 1.$$

Therefore, using the general equation $\|\mathsf{q}\| = q_R{}^2 + \|\mathbf{q}_P\|^2$ we know that

$$q_R{}^2 + \|\mathbf{q}_P\|^2 = 1 \Rightarrow q_R{}^2 - 1 = -\|\mathsf{q}\|^2 \Rightarrow 2q_R{}^2 - 1 = q_R{}^2 - \|\mathsf{q}\|^2 .$$

The RHS of this resulting equation is the real part of Equation (4). Plugging this into that equation shows us that

$$\mathsf{q}^2 = 2q_R{}^2 - 1 + (2q_R\mathbf{q}_P) = 2q_R \cdot \left(q_R + \overbrace{\mathbf{q}_P}^{\mathsf{q}}\right) - 1 = 2q_R \cdot \mathsf{q} - 1$$

as desired.

(2) Just take the real part of the resulting equation. Note that this matches the equation:

$$\cos 2\theta = 2\cos^2\theta - 1.$$

3.46 (3.46) Let q denote the quaternion, $\frac{2}{5} + \frac{1}{5}\mathbf{i} + \frac{2}{5}\mathbf{j} - \frac{4}{5}\mathbf{k}$, in question.

- Since

$$\left(\frac{2}{5}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(\frac{2}{5}\right)^2 + \left(\frac{4}{5}\right)^2 = \frac{4}{25} + \frac{1}{25} + \frac{4}{25} + \frac{16}{25} = \frac{25}{25} = 1$$

  we see that $\|\mathsf{q}\| = 1$ and as such is an element of $\mathbb{S}^3$.

- As an element of $\mathbb{S}^3$ we know that it's inverse is its conjugate: that is

$$\mathsf{q}^{-1} = \overline{\mathsf{q}} = \frac{2}{5} - \frac{1}{5}\mathbf{i} - \frac{2}{5}\mathbf{j} + \frac{4}{5}\mathbf{k}.$$

- By Exercise 3.45, we know that

$$\begin{aligned}
\mathsf{q}^2 &= \left(2 \cdot \frac{2}{5}\right)\mathsf{q} - 1 \\
&= \left(2 \cdot \frac{2}{5}\right)\left(\frac{2}{5} + \frac{1}{5}\mathbf{i} + \frac{2}{5}\mathbf{j} - \frac{4}{5}\mathbf{k}\right) - 1 \\
&= \left(\frac{8}{25} + \frac{4}{25}\mathbf{i} + \frac{8}{25}\mathbf{j} - \frac{16}{25}\mathbf{k}\right) - 1 \\
&= -\frac{17}{25} + \frac{4}{25}\mathbf{i} + \frac{8}{25}\mathbf{j} - \frac{16}{25}.
\end{aligned}$$

3.48 (3.48) In this problem we are asked to find the powers of the quaternion

$$\alpha \stackrel{\text{def}}{=} \frac{\sqrt{3}}{2} + \frac{3}{11}\mathbf{i} + \frac{1}{11}\mathbf{j} - \frac{9}{22}\mathbf{k}.$$

For starters, we have that

$$\|\alpha\|^2 = \frac{3}{4} + \frac{9}{121} + \frac{1}{121} + \frac{81}{4 \cdot 121} = 1.$$

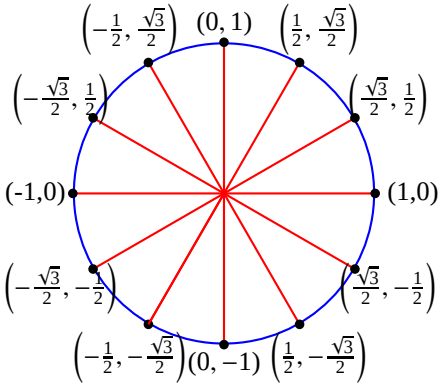


Figure L.1: Figure 52

Now Figure 52 is the unit circle we're interested in. Therefore there are 12 distinct powers of α.

- The zero power of α is just 1.
- Of course the first power of α is

  $$\alpha = \frac{\sqrt{3}}{2} + \frac{3}{11}\mathbf{i} + \frac{1}{11}\mathbf{j} - \frac{9}{22}\mathbf{k}.$$

  itself.
- Looking at the unit circle, the second power has coordinates $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$. So to get α^2 we need to divide the pure part by $\sqrt{3}$ and multiply the pure part by $\sqrt{3}$. Thus

  $$\alpha^2 = \frac{1}{2} + \frac{3\sqrt{3}}{11}\mathbf{i} + \frac{\sqrt{3}}{11}\mathbf{j} - \frac{9\sqrt{3}}{22}\mathbf{k}.$$

- The coordinates for α^3 are $(0, 1)$. Therefore we need to ignore the real part of α and multiply the pure part of α by 2. Therefore

  $$\alpha^3 = \frac{6}{11}\mathbf{i} + \frac{2}{11}\mathbf{j} - \frac{9}{11}\mathbf{k}.$$

- From the circle we can see that α^4 is just the negative conjugate of α^2 (i.e. we flip the sign of the real part and leave alone its pure part). Therefore
$$\alpha^4 = \frac{-1}{2} + \frac{3\sqrt{3}}{11}\mathbf{i} + \frac{\sqrt{3}}{11}\mathbf{j} - \frac{9\sqrt{3}}{22}\mathbf{k}.$$
- From the circle we can see that α^5 is just the negative conjugate of α (i.e. we flip the sign of the real part and leave alone its pure part). Therefore
$$\alpha^5 = -\frac{\sqrt{3}}{2} + \frac{3}{11}\mathbf{i} + \frac{1}{11}\mathbf{j} - \frac{9}{22}\mathbf{k}.$$
- From the circle we can see that α^6 is just the negative conjugate of α^0 (i.e. we flip the sign of the real part and leave alone its pure part). Therefore
$$\alpha^6 = -1.$$

We can now easily compute the other 6 powers since they are all the negatives of one of the first six powers. Therefore we have

- $\alpha^7 = \alpha^6 * \alpha = -\alpha = -\frac{\sqrt{3}}{2} - \frac{3}{11}\mathbf{i} - \frac{1}{11}\mathbf{j} + \frac{9}{22}\mathbf{k}$
- $\alpha^8 = -\alpha^2 = -\frac{1}{2} - \frac{3\sqrt{3}}{11}\mathbf{i} - \frac{\sqrt{3}}{11}\mathbf{j} + \frac{9\sqrt{3}}{22}\mathbf{k}.$
- $\alpha^9 = -\alpha^3 = -\frac{6}{11}\mathbf{i} - \frac{2}{11}\mathbf{j} + \frac{9}{11}\mathbf{k}.$
- $\alpha^{10} = -\alpha^4 = \overline{\alpha^2} = \frac{1}{2} - \frac{3\sqrt{3}}{11}\mathbf{i} - \frac{\sqrt{3}}{11}\mathbf{j} + \frac{9\sqrt{3}}{22}\mathbf{k}.$
- $\alpha^{11} = -\alpha^5 = \overline{\alpha} = \frac{\sqrt{3}}{2} - \frac{3}{11}\mathbf{i} - \frac{1}{11}\mathbf{j} + \frac{9}{22}\mathbf{k}.$

Lastly, the question asks us to find α^{117}. But since $117 = 9 \cdot 12 + 9$ we have that
$$\alpha^{117} = \alpha^9 = -\alpha^3 = -\frac{6}{11}\mathbf{i} - \frac{2}{11}\mathbf{j} + \frac{9}{11}\mathbf{k}.$$

3.50 (3.50) The difference of two squares property $x^2 - y^2 = (x - y)(x + y)$ depends on commutativity. Indeed expanding the RHS we have that
$$(x - y)(x + y) = x^2 - yx + xy - y^2$$
which equals $x^2 - y^2$ iff $-yx + xy = 0$, i.e. $xy = yx$.

3.51 (3.51)
$$\begin{aligned}
S_{\widehat{\mathbf{u}}} * e^{\theta,\widehat{\mathbf{v}}} * S_{\widehat{\mathbf{u}}}^{-1} &= \cos\theta + \sin\theta S_{\widehat{\mathbf{u}}} * \widehat{\mathbf{v}} * S_{\widehat{\mathbf{u}}}^{-1} \\
&= \cos\theta + \sin\theta S_{\widehat{\mathbf{u}}} * S_{\widehat{\mathbf{u}}} * \widehat{\mathbf{v}} && (\{\text{give ref}\}) \\
&= \cos\theta + \sin\theta S_{\widehat{\mathbf{u}}}^2 * \widehat{\mathbf{v}} \\
&= \cos\theta + \sin\theta \widehat{\mathbf{u}} * \widehat{\mathbf{v}} \\
&= \cos\theta + \sin\theta \widehat{\mathbf{u}} \times \widehat{\mathbf{v}} && (\text{Quat mul using } \widehat{\mathbf{u}} \perp \widehat{\mathbf{u}}') \\
&= e^{\theta,\widehat{\mathbf{u}}\times\widehat{\mathbf{v}}}.
\end{aligned}$$

3.54 (3.54)

⇒ Suppose that $R^\times$ is closed under multiplication. Then for all r and $s \in R^\times$ we have $r \cdot s \in R^\times$. Translating this, this means for all r and s not equal to 0, $r \cdot s \neq 0$. But this is exactly the same as the statement that R contains not zero divisors.

⇐ All statements can be turned around.

3.57 (3.57)

(1) ($[\,5\,]_{21}$) Since $5 \cdot 4 = 20 \equiv_{21} -1$ we know that

$$5 \cdot (-4) \equiv_{21} 1 \Rightarrow [\,5\,]_{21}^{-1} = [\,-4\,]_{21} = [\,17\,]_{21}\,.$$

(2) ($[\,4\,]_{17}$) Since $4^2 = 16 \equiv_{17} -1$ we know that

$$[\,4\,]_{17}^{-1} = [\,-4\,]_{17} = [\,13\,]_{17}\,.$$

(3) ($[\,7\,]_{11}$) Since $7 \cdot 8 = 56 \equiv_{11} 1$ we see directly that $[\,7\,]_{11}^{-1} = [\,8\,]_{11}$.

(4) ($[\,21\,]_{35}$) Since 21 and 35 are not relatively prime (the number 7 divides both), $[\,21\,]_{35}$ has no inverse.

3.58 (3.58)

3.60 (3.60)

3.61 (3.61)

3.62 (3.62)

3.63 (3.63)

3.68 (3.68)

3.79 (3.79) Suppose that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are linearly dependent. Denote the determinant in question as D. Then, by definition, there exist scalars $r_1, \ldots, r_n \in R$, not all zero, so that

$$\mathbf{0} = r_1\mathbf{v}_1 + \cdots r_n\mathbf{v}_n.$$

By renumbering, if necessary, we can assume that $r_1 \neq 0$. Then we have that

$$r_1\mathbf{v}_1 = -(r_2\mathbf{v}_2 + \cdots r_n\mathbf{v}_n).$$

Therefore, we have that

$$\begin{aligned}
r_1 \cdot \mathrm{D} &= \det(\mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n) \\
&= \det(r_1 \cdot \mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n) \\
&= -\det(r_2\mathbf{v}_2 + \cdots r_n\mathbf{v}_n \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n) && \text{(Replace } r_1 \cdot \mathbf{v}_1\text{)} \\
&= -\det(r_2\mathbf{v}_2 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n) - \cdots - \det(r_n\mathbf{v}_n \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n) && \text{(Dist)} \\
&= -r_2 \det(\mathbf{v}_2 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n) - \cdots - r_n \det(\mathbf{v}_n \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n) && \text{(Scalar pull)} \\
&= -r_2 \cdot 0 - \cdots - r_n \cdot 0 && \text{(Det is 0 whenever 2 arg are same)} \\
&= 0.
\end{aligned}$$

Since $r_1 \neq 0$, we know that D is a zero divisor.

3.88 (3.88) For any $\mathbf{A}, \mathbf{B} \in \mathbf{M}_n(R)$, we have

$$\begin{aligned}\operatorname{tr}(\mathbf{A}+\mathbf{B}) &= \sum_{i=1}^{n}(\mathbf{A}+\mathbf{B})(i,i)\\ &= \sum_{i=1}^{n}(\mathbf{A}(i,i)+\mathbf{B}(i,i))\\ &= \sum_{i=1}^{n}\mathbf{A}(i,i)+\sum_{i=1}^{n}\mathbf{B}(i,i)\\ &= \operatorname{tr}(\mathbf{A})+\operatorname{tr}(\mathbf{B}).\end{aligned}$$

Therefore tr preserves addition. Meanwhile we have for any $r \in R$

$$\begin{aligned}\operatorname{tr}(r\cdot\mathbf{A}) &= \sum_{i=1}^{n}(r\cdot\mathbf{A})(i,i)\\ &= \sum_{i=1}^{n} r\cdot(\mathbf{A}(i,i))\\ &= r\cdot\sum_{i=1}^{n}\mathbf{A}(i,i)\\ &= r\cdot\operatorname{tr}(\mathbf{A}).\end{aligned}$$

Therefore tr preserves scalar multiplication.

3.90 (3.90) As pointed out in the hint, you could simply do this by doing a direct calculation. If you cheat and use results from the next section about the adjoint, a slicker proof can be obtained by noting that

$$\mathbf{A}+\widetilde{\mathbf{A}} = \operatorname{tr}(\mathbf{A})\cdot\mathbf{I}_2.$$

Then multiply both sides of this equation by $\mathbf{A}$ to get

$$\mathbf{A}^2+\mathbf{A}\cdot\widetilde{\mathbf{A}} = \operatorname{tr}(\mathbf{A})\cdot\mathbf{A} = \mathbf{A}^2+\det\mathbf{A}\cdot\mathbf{I}_2 = \operatorname{tr}(\mathbf{A})\cdot\mathbf{A}.$$

and then rearrange to finish the exercise.

4.1 (4.1)

(1) $(\mathbb{N},+)$
Yes

(2) $(\mathbb{N},-)$
No $1-2\notin\mathbb{N}$.

(3) $(\mathbb{Z},+)$
Yes

(4) $(\mathbb{Z},\div)$
No For two reasons. The first is that $1/0$ is not defined (so that $(1,0)$ is not in the domain of $\div$) the other is that $1/2$ is not an integer.

(5) $(\mathbb{Z}^\times, \div)$
No The second of the two problems still occurs here: i.e. $1/2$ is not an integer.

(6) $(\mathbb{Q}, \div)$
No Division by 0 again is the problem.

(7) $(\mathbb{Q}^\times, \div)$
Yes

(8) $(\mathcal{P}(X) \setminus \{\varnothing\}, \cup)$
Yes The union of two non-empty sets is non-empty.

(9) $(\mathcal{P}(X), \setminus)$
Yes Yes, set difference of two subsets of X is a subset of X.

(10) $(\mathcal{P}(X) \setminus \{\varnothing\}, \setminus)$
No (Unless $X = \varnothing$) Set difference between any set and itself is empty.

(11) $(\{\{5\}\}, \cdot)$
Yes Product of two roots of unity is a root of unity.

(12) $(\{\{5\}\}, +)$
No $1 \in \{\{5\}\}$ but $1 + 1 \notin \{\{5\}\}$.

(13) $(\mathcal{F}_X, \circ)$
Yes Composition of two functions $X \to X$ is still a function $X \to X$.

(14) $(\{0, 1\}, \cdot)$
Yes Since all four combinations of multiplication of 1 and 0 gives either 1 or 0.

(15) $(\{0, 1\}, +)$
No $1 + 1 = 2$ which is not either 0 or 1.

(16) $(\mathbb{S}^3, *)$
Yes This was just mentioned in Example 4.7 back on page 215.

(17) $(\mathbb{R}^3, \times)$
Yes The cross product of two vectors is a vector.

(18) $(\mathbb{R}^3, \bullet)$
No The dot product of two vectors is a number, not a vector.

(19) $(\mathbb{Z}_5, \cdot)$
Yes Yes, since 5 product of two non-zero integers $\mod 5$ is a non-zero integer $\mod 5$.

(20) $(\mathbb{Z}_6^\times, \cdot)$
No $[\,2\,]_6 \cdot [\,3\,]_6 = [\,0\,]_6$ and so $\cdot$ is not closed on $\mathbb{Z}_6^\times$.

(21) $(\mathcal{F}\mathrm{un}\,(X, Y)\,, \circ)$
No Two functions $X \to Y$ cannot be composed (unless $X = Y$).

4.5 (4.5)

4.12 (4.12)

4.15 (4.15)

4.16 (4.16)

(1) The number 1 is an identity for $(\mathbb{H}, \divideontimes)$.

(2)

$$\mathbf{i} \divideontimes \mathbf{i} = 1 \Rightarrow (\mathbf{i} \divideontimes \mathbf{i}) \divideontimes \mathbf{j} = 1 \divideontimes \mathbf{j} = \mathbf{j}$$

but

$$\mathbf{i} \divideontimes \mathbf{j} = \mathbf{k} \Rightarrow \mathbf{i} \divideontimes (\mathbf{i} \divideontimes \mathbf{j}) = \mathbf{i} \divideontimes \mathbf{k} = -\mathbf{j}.$$

Since $\mathbf{j} \neq -\mathbf{j}$, the $\divideontimes$ is not associative.

(3) Let q, q', and $\mathsf{q}'' \in \mathbb{H}$ be arbitrary. Then, we have

$$\begin{aligned}
\mathbf{Re}((\mathsf{q} \divideontimes \mathsf{q}') \divideontimes \mathsf{q}'') &= \mathbf{Re}((q_R \cdot q'_R + \mathbf{q}_P \bullet \mathbf{q}'_P, q_R \cdot \mathbf{q}'_P + q'_R \cdot \mathbf{q}_P + \mathbf{q}_P \times \mathbf{q}'_P) \divideontimes \mathsf{q}'') \\
&= (q_R \cdot q'_R + \mathbf{q}_P \bullet \mathbf{q}'_P) \cdot q''_R + (q_R \cdot \mathbf{q}'_P + q'_R \cdot \mathbf{q}_P + \mathbf{q}_P \times \mathbf{q}'_P) \bullet \mathbf{q}''_P \\
&= q_R \cdot q'_R \cdot q''_R + q_R(\mathbf{q}'_P \bullet \mathbf{q}''_P) + q'_R(\mathbf{q}''_P \bullet \mathbf{q}_P) + q''_R(\mathbf{q}_P \bullet \mathbf{q}'_P) + (\mathbf{q}_P \times \mathbf{q}'_P) \bullet \mathbf{q}''_P.
\end{aligned}$$

On the other hand, the same sort of calculation shows us that

$$\mathbf{Re}(\mathsf{q} \divideontimes (\mathsf{q}' \divideontimes \mathsf{q}'')) = q_R \cdot q'_R \cdot q''_R + q_R(\mathbf{q}'_P \bullet \mathbf{q}''_P) + q'_R(\mathbf{q}''_P \bullet \mathbf{q}_P) + q''_R(\mathbf{q}_P \bullet \mathbf{q}'_P) + \mathbf{q}_P \bullet (\mathbf{q}'_P \times \mathbf{q}''_P).$$

The first four terms of these two expressions are identical. And we are done since we also have equality of the last term in each of the expressions since

$$(\mathbf{q}_P \times \mathbf{q}'_P) \bullet \mathbf{q}''_P = \det(\mathbf{q}_P, \mathbf{q}'_P, \mathbf{q}''_P) = \mathbf{q}_P \bullet (\mathbf{q}'_P \times \mathbf{q}''_P).$$

(4) A direct calculation shows that in general $\mathsf{q} \divideontimes \overline{\mathsf{q}} = q_R^2 - \| \mathbf{q}_P \|^2$. Therefore we always have that $\mathsf{q} \divideontimes \overline{\mathsf{q}}$ is a real number which equals zero iff q is on the light cone. This proves two things:

- If q is not on the light cone,

$$\frac{\mathsf{q}}{q_R^2 - \| \mathbf{q}_P \|^2}$$

  is an inverse to q.

- If q is on the light cone, then q is a zero divisor.[III] Now we assume that q is on the light cone and has an inverse p. Then, we have that

$$\mathsf{q} \divideontimes \mathsf{p} = 1 \Rightarrow \overline{\mathsf{q}} \divideontimes (\mathsf{q} \divideontimes \mathsf{p}) = \overline{\mathsf{q}}.$$

[III] If we were in an associative setting we'd be done at this very moment!

But since $\divideontimes$ is real part associative, we have that

$$\begin{aligned}\mathbf{Re}\, q &= \mathbf{Re}\, \overline{q} \\ &= \mathbf{Re}(\overline{q} \divideontimes (q \divideontimes p)) \\ &= \mathbf{Re}((\overline{q} \divideontimes q) \divideontimes p) \\ &= \mathbf{Re}(\mathbf{0} \divideontimes p) \\ &= \mathbf{Re}(\mathbf{0}) \\ &= 0.\end{aligned}$$

This implies that q is pure. But for pure quaternions we have that

$$q \divideontimes q = \| q \|^2 \Rightarrow \overline{q} \divideontimes q = -\| q \|^2 \Rightarrow \| q \| = 0.$$

But since $\| q \| > 0$ for all non-zero elements, we know that $q = \mathbf{0}$.

5.3 (5.3)

(1)

(2)

(3)

(4)

(5)

(6)

5.14 (5.14)

$(+, +)$	$([0]_3, [0]_2)$	$([1]_3, [0]_2)$	$([2]_3, [0]_2)$	$([0]_3, [1]_2)$	$([1]_3, [1]_2)$	$([2]_3, [0]_2)$
$([0]_3, [0]_2)$	$([0]_3, [0]_2)$	$([1]_3, [0]_2)$	$([2]_3, [0]_2)$	$([0]_3, [1]_2)$	$([1]_3, [1]_2)$	$([2]_3, [0]_2)$
$([1]_3, [0]_2)$	$([1]_3, [0]_2)$	$([2]_3, [0]_2)$	$([0]_3, [0]_2)$	$([1]_3, [1]_2)$	$([2]_3, [1]_2)$	$([0]_3, [0]_2)$
$([2]_3, [0]_2)$	$([2]_3, [0]_2)$	$([0]_3, [0]_2)$	$([1]_3, [0]_2)$	$([2]_3, [1]_2)$	$([0]_3, [1]_2)$	$([1]_3, [0]_2)$
$([0]_3, [1]_2)$	$([0]_3, [1]_2)$	$([1]_3, [1]_2)$	$([2]_3, [1]_2)$	$([0]_3, [0]_2)$	$([1]_3, [0]_2)$	$([2]_3, [1]_2)$
$([1]_3, [1]_2)$	$([1]_3, [1]_2)$	$([2]_3, [1]_2)$	$([0]_3, [1]_2)$	$([1]_3, [0]_2)$	$([2]_3, [0]_2)$	$([0]_3, [1]_2)$
$([1]_3, [1]_2)$	$([2]_3, [1]_2)$	$([0]_3, [1]_2)$	$([1]_3, [1]_2)$	$([2]_3, [0]_2)$	$([0]_3, [0]_2)$	$([1]_3, [1]_2)$

5.15 (5.15) Since this table is long we write elements of $\mathbb{Z}_2$ without brackets.

$(+,+)$	$(\mathbf{1},0)$	$(\rho,0)$	$(\rho^2,0)$	$(\mathbf{r},0)$	$(\mathbf{r}\rho,0)$	$(\mathbf{r}\rho^2,0)$	$(\mathbf{1},1)$	$(\rho,1)$	$(\rho^2,1)$	$(\mathbf{r},1)$	$(\mathbf{r}\rho,1)$	$(\mathbf{r}\rho^2,1)$
$(\mathbf{1},0)$	$(\mathbf{1},0)$	$(\rho,0)$	$(\rho^2,0)$	$(\mathbf{r},0)$	$(\mathbf{r}\rho,0)$	$(\mathbf{r}\rho^2,0)$	$(\mathbf{1},1)$	$(\rho,1)$	$(\rho^2,1)$	$(\mathbf{r},1)$	$(\mathbf{r}\rho,1)$	$(\mathbf{r}\rho^2,1)$
$(\rho,0)$	$(\rho,0)$	$(\rho^2,0)$	$(\mathbf{1},0)$	$(\mathbf{r}\rho^2,0)$	$(\mathbf{r},0)$	$(\mathbf{r}\rho,0)$	$(\rho,1)$	$(\rho^2,1)$	$(\mathbf{1},1)$	$(\mathbf{r}\rho^2,1)$	$(\mathbf{r},1)$	$(\mathbf{r}\rho,1)$
$(\rho^2,0)$	$(\rho^2,0)$	$(\mathbf{1},0)$	$(\rho,0)$	$(\mathbf{r}\rho,0)$	$(\mathbf{r}\rho^2,0)$	$(\mathbf{r},0)$	$(\rho^2,1)$	$(\mathbf{1},1)$	$(\rho,1)$	$(\mathbf{r}\rho,1)$	$(\mathbf{r}\rho^2,1)$	$(\mathbf{r},1)$
$(\mathbf{r},0)$	$(\mathbf{r},0)$	$(\mathbf{r}\rho,0)$	$(\mathbf{r}\rho^2,0)$	$(\mathbf{1},0)$	$(\rho,0)$	$(\rho^2,0)$	$(\mathbf{r},1)$	$(\rho^2\mathbf{r},1)$	$(\rho\mathbf{r},1)$	$(\mathbf{1},1)$	$(\rho,1)$	$(\rho^2,0)$
$(\mathbf{r}\rho,0)$	$(\mathbf{r}\rho,0)$	$(\mathbf{r}\rho^2,0)$	$(\mathbf{r},0)$	$(\rho^2,0)$	$(\mathbf{1},0)$	$(\rho,0)$	$(\mathbf{r}\rho,1)$	$(\mathbf{r}\rho^2,1)$	$(\mathbf{r},1)$	$(\rho^2,1)$	$(\mathbf{1},1)$	$(\rho,1)$
$(\mathbf{r}\rho^2,0)$	$(\mathbf{r}\rho^2,0)$	$(\mathbf{r},0)$	$(\mathbf{r}\rho,0)$	$(\rho,0)$	$(\rho^2,0)$	$(\mathbf{1},0)$	$(\mathbf{r}\rho^2,1)$	$(\mathbf{r},1)$	$(\mathbf{r}\rho,1)$	$(\rho,1)$	$(\rho^2,1)$	$(\mathbf{1},1)$
$(\mathbf{1},1)$	$(\mathbf{1},1)$	$(\rho,1)$	$(\rho^2,1)$	$(\mathbf{r},1)$	$(\mathbf{r}\rho,1)$	$(\mathbf{r}\rho^2,1)$	$(\mathbf{1},0)$	$(\rho,0)$	$(\rho^2,0)$	$(\mathbf{r},0)$	$(\mathbf{r}\rho,0)$	$(\mathbf{r}\rho^2,0)$
$(\rho,1)$	$(\rho,1)$	$(\rho^2,1)$	$(\mathbf{1},1)$	$(\mathbf{r}\rho^2,1)$	$(\mathbf{r},1)$	$(\mathbf{r}\rho,1)$	$(\rho,0)$	$(\rho^2,0)$	$(\mathbf{1},0)$	$(\mathbf{r}\rho^2,0)$	$(\mathbf{r},0)$	$(\mathbf{r}\rho,0)$
$(\rho^2,1)$	$(\rho^2,1)$	$(\mathbf{1},1)$	$(\rho,1)$	$(\mathbf{r}\rho,1)$	$(\mathbf{r}\rho^2,1)$	$(\mathbf{r},1)$	$(\rho^2,0)$	$(\mathbf{1},0)$	$(\rho,0)$	$(\mathbf{r}\rho,0)$	$(\mathbf{r}\rho^2,0)$	$(\mathbf{r},0)$
$(\mathbf{r},1)$	$(\mathbf{r},1)$	$(\mathbf{r}\rho,1)$	$(\mathbf{r}\rho^2,1)$	$(\mathbf{1},1)$	$(\rho,1)$	$(\rho^2,1)$	$(\mathbf{r},0)$	$(\rho^2\mathbf{r},0)$	$(\rho\mathbf{r},0)$	$(\mathbf{1},0)$	$(\rho,0)$	$(\rho^2,1)$
$(\mathbf{r}\rho,1)$	$(\mathbf{r}\rho,1)$	$(\mathbf{r}\rho^2,1)$	$(\mathbf{r},1)$	$(\rho^2,1)$	$(\mathbf{1},1)$	$(\rho,1)$	$(\mathbf{r}\rho,0)$	$(\mathbf{r}\rho^2,0)$	$(\mathbf{r},0)$	$(\rho^2,0)$	$(\mathbf{1},0)$	$(\rho,0)$
$(\mathbf{r}\rho^2,1)$	$(\mathbf{r}\rho^2,1)$	$(\mathbf{r},1)$	$(\mathbf{r}\rho,1)$	$(\rho,1)$	$(\rho^2,1)$	$(\mathbf{1},1)$	$(\mathbf{r}\rho^2,0)$	$(\mathbf{r},0)$	$(\mathbf{r}\rho,0)$	$(\rho,0)$	$(\rho^2,0)$	$(\mathbf{1},0)$

5.17 (5.17) It is "the same as" the symmetries of the rectangle since every symmetry of the letter 'H' is an involution (meaning if you do it twice you've done nothing).

5.19 (5.19)

5.20 (5.20)

5.24 (5.24)

5.25 (5.25)

5.26 (5.26) To make the use of Theorem 5.38 more explicit, we replace the letter g by the letter h. So we are trying to prove that

$$|h^r| \mid |h|.$$

Having freed up the symbol g, we use it to denote instead h^r (so we've ended up replacing the original g^r with g). We also denote $|g|$ as n and $|h|$ as m. So we wish to prove that

$$n \mid m.$$

With these notational changes, it is now clear how we should use Theorem 5.38**(??)**:

$$\begin{aligned} g^m &= (h^r)^m \\ &= h^{rm} \\ &= h^{mr} \\ &= (h^m)^r \\ &= (e_G)^r \quad (|h| = m) \\ &e_G. \end{aligned}$$

So indeed, Theorem 5.38**(??)** to conclude that $m \mid n$ as desired.

5.27 (5.27) Let $d := \gcd(n, m)$. I give two proofs of this exercise. Of course you only need to provide one.

(1) Using Equation (2) we see that $G[\,n\,] = \bigcup_{d|n} G(d)$ and $G[\,m\,] = \bigcup_{d|m} G(d)$. Now using Equation (3) and the basic properties of intersection an union, we have

$$G[\,n\,] \cap G[\,m\,] = \bigcup_{d|n \text{ and } d|m} G(d).$$

But the set $\{d \in \mathbb{Z} : d \mid n \text{ and } d \mid m\}$ is the same as the set of divisors of $\gcd(n, m)$, essentially by the very definition of the greatest common divisor. Therefore

$$G[\,n\,] \cap G[\,m\,] = \bigcup_{d|\gcd(n,m)} G(d) = G[\,\gcd(n, m)\,].$$

(2)$\subseteq$ Let d By the Euclidean algorithm formula there exist integers a and b so that

$$d = na + mb.$$

Suppose

$$\begin{aligned} g \in G[\,n\,] \cap G[\,m\,] &\Rightarrow g \in G[\,n\,] \text{ and } g \in G[\,m\,] \\ &\Rightarrow g^n = e_G \text{ and } g^m = e_G \\ &\Rightarrow g^{na} = (g^n)^a = e_G^a = e_G \text{ and } g^{mb} = (g^m)^b = (e_G)^b = e_G \\ &\Rightarrow g^d = g^{na+mb} = g^{na} g^{mb} = e_G \cdot e_G = e_G \\ &\Rightarrow g \in G[\,d\,]. \end{aligned}$$

$\supseteq$ Conversely, suppose that $g \in G[\,d\,]$. Then, since $d \mid n$ and $d \mid m$ we know that $g^n = e_G$ and $g^m = e_G$. Therefore $g \in G[\,n\,] \cap G[\,m\,]$ as promised.

5.30 (5.30)

5.43 (5.43)

5.47 (5.47)

5.48 (5.48) Since $4 \cdot ([\,a\,]_4, [\,b\,]_4) = ([\,0\,]_4, [\,0\,]_4)$ every element of $\mathbb{Z}_4$ has an order which divides 4, i.e must have order 1, 2, or 4. An element $([\,a\,]_4, [\,b\,]_4) \in \mathbb{Z}_4^2$ has order 2 iff it is not the identity and we have

$$2 \cdot ([\,a\,]_4, [\,b\,]_4) = ([\,0\,]_4, [\,0\,]_4).$$

In other words, iff $[\,a\,]_4$ and $[\,b\,]_4$ are in $\mathbb{Z}_4[2]$, which has $\gcd(2, 4) = 2$ elements, and are not both $[\,0\,]_4$. Therefore there are 3 elements of order 2 in $\mathbb{Z}_4^2[2]$, they are

$$([\,2\,]_4, [\,0\,]_4) \qquad ([\,0\,]_4, [\,2\,]_4) \qquad ([\,2\,]_4, [\,2\,]_4).$$

After the identity and these 3 elements, the remaining $16 - 3 - 1 = 12$ elements of $\mathbb{Z}_4^2$ must have order 4 (that's the only possibility remaining). In summary we have the following order table: (Discussed further in Definition 8.57. It's just an organized account of our results in this exercise.)

$\mathbb{Z}_4^2$	n	1	2	4
	$\mathbf{o}_n$	1	3	12

5.49 (5.49) We know that (omitting mention of $\mathcal{S}_3 \times \mathbb{Z}_3$ in these calculations)

$$\mathbf{O}_{18} = 6 \cdot 3 = 18 \qquad \mathbf{O}_9 = 3 \cdot 3 = 9 \qquad \mathbf{O}_6 = 6 \cdot 3 = 18$$
$$\mathbf{O}_3 = 3 \cdot 3 = 9 \qquad \mathbf{O}_2 = 4 \cdot 1 = 4 \qquad \mathbf{O}_1 = 1 \cdot 1 = 1.$$

Therefore

$$\mathbf{o}_1 = 1 \qquad \mathbf{o}_2 = \mathbf{O}_2 - 1 = 4 - 1 = 3$$
$$\mathbf{o}_3 = \mathbf{O}_3 - 1 = 9 - 1 = 8 \qquad \mathbf{o}_6 = \mathbf{O}_6 - \mathbf{O}_3 - \mathbf{O}_2 + \mathbf{O}_1 = 6,$$

and we've found all $1 + 3 + 8 + 6 = 18$ elements of $\mathcal{S}_3 \times \mathbb{Z}_3$

$\mathcal{S}_3 \times \mathbb{Z}_3$	n	1	2	3	6
	$\mathbf{o}_n$	1	3	8	6

.

Actually, this example is small enough that we didn't need to use the formula. For instance, the only elements of order 6 are of the form

$$(\tau, a)$$

where τ is a transposition and a is a non-identity element of $\mathbb{Z}_3$. There are 3 transpositions and 2 non-identity elements of $\mathbb{Z}_3$, leading to the 6 elements order order 6.

2 (5.51)

(1)

(2) Consider the set of integers $\{6, 10, 15\}$. Then we have

$$\gcd(6, 10) = 2 \qquad \gcd(10, 15) = 5 \qquad \gcd(6, 15) = 3$$

so no pair consists of relatively prime integers. But

$$\gcd(6, 10, 15) = \gcd(\gcd(6, 10), 15) = \gcd(2, 15) = 1.$$

5.52 (5.52)

5.53 (5.53)

5.58 (5.58)

(1) Let $g \in G$. Then $g = x^a$ for some integer a. Then $g^m = (x^a)^m = (x^m)^a$. Therefore $g^m \in \langle\, x^m \,\rangle$. Since an arbitrary element of $C\hat{}\,m$ has the form we are done.

(2) Since $C\hat{}\,m = \langle\, x^m \,\rangle$ by the first part, and since $\langle\, x^m \,\rangle = \left\langle\, x^{\gcd(n,m)} \,\right\rangle$ we have $C\hat{}\,m = C\hat{}\,\gcd(n,m)$. Therefore, we need to prove that (we can flip the divisor over for notational ease)

$$d \mid n \Rightarrow C\hat{}\left[\frac{n}{d}\right] = C[d].$$

$\subseteq$ One direction is easy: if $h \in C\hat{}\left[\frac{n}{d}\right]$, then $h = g^{\frac{n}{d}}$ for some $g \in C$. Therefore, $(g^{\frac{n}{d}})^d = g^n = e_C$ and $h = g^{\frac{n}{d}} \in C[d]$.

$\supseteq$ On the other hand, if $h \in C[d]$, then first of all since $h \in \langle\, x \,\rangle$, there is some integer a so that $h = x^a$, where x is a generator of C, and second of all we also know that know that $x^{da} = e_C$. But this means that $n \mid da$, i.e. there exists some integer q so that $da = qn$, or $a = \frac{qn}{d} = q \cdot \frac{n}{d}$. Therefore

$$h = x^a = x^{q \cdot \frac{n}{d}} = x^{q\frac{n}{d}} \in C\hat{}\,[n/d]$$

as desired.

5.62 (5.62)

(1) In the group $\mathcal{D}_4$, the only elements which do not equal their own inverse are ρ and ρ^{-1}. However, these two elements are conjugate to each other via $\mathbf{r}$.

(2) I present three possibilities:

(a) In the group $B\mathcal{D}_3$ the inverse of the $\mathbf{j}$ is $-\mathbf{j}$. Since every element of $B\mathcal{D}_3$ has the form $\chi_6^m \cdot \mathbf{j}^n$, we have that

$$\{\mathfrak{q}\mathbf{j}\mathfrak{q}^{-1} : \mathfrak{q} \in B\mathcal{D}_3\} = \{\chi_6^m \cdot \mathbf{j}^n\mathbf{j}\mathbf{j}^{-n} \cdot \chi_6^{-m} : m \in \mathbb{Z}\} = \{\chi_6^m \cdot \mathbf{j} \cdot \chi_6^m : m \in \mathbb{Z}\}.$$

Moreover,

$$\chi_6^m \mathbf{j} \chi_6^{-m} = \chi_6^m \chi_6^m \mathbf{j} = \chi_6^{2m}\mathbf{j} = \chi_3^m \mathbf{j}.$$

But as 3 is odd, no power of $\chi_3 = -1$ (as -1 is not a cube root of unity– $(-1)^3 = -1 \neq 1$). Therefore, just a conjugation cannot equal $-\mathbf{j}$.

(b) Another example where this happens is $\mathbf{SL}_2(\mathbb{Z}_3)$. There the matrices

$$\mathfrak{Q} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathfrak{Q}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

are inverse to each other. However, if they were conjugate via $\mathbf{A} \in \mathbf{SL}_2(\mathbb{Z}_3)$, then

$$\mathbf{A}\mathfrak{Q}\mathbf{A}^{-1} = \mathfrak{Q}^{-1} \Rightarrow \mathfrak{Q}\mathbf{A}\mathfrak{Q} = \mathbf{A}.$$

Writing in $\mathbf{A}$ as

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow \mathfrak{Q}\mathbf{A} = \begin{pmatrix} a+c & b+d \\ c & d \end{pmatrix} \Rightarrow \mathfrak{Q}\mathbf{A}\mathfrak{Q} = \begin{pmatrix} a+c & a+b+c+d \\ c & c+d \end{pmatrix}.$$

Setting this equal to $\mathbf{A}$ implies that $c = 0$ and that $a + d$ also equals zero, i.e. $d = -a$. So

$$\begin{pmatrix} a & b \\ 0 & -a \end{pmatrix}.$$

Lastly, since $\mathbf{A}$ is to have determinant 1 we need to have $a^2 = -1$. But this equation is not solvable in $\mathbb{Z}_3$. Actually, this proves that $\mathfrak{Q}$ and $\mathfrak{Q}$ are conjugate in $\mathbf{SL}_2\left(\mathbb{Z}_p\right)$ iff $p \equiv_4 1$.

(c) The most elementary example where this happens is $\mathcal{A}_4$, but seeing we've yet to actually prove that $\mathcal{A}_4$ is a group this might not be the best example to give. In terms of future notation, one can show that the inverse pair of permutations $(1\ 2\ 3)$ and $(1\ 3\ 2)$ are not conjugate.

5.72 (5.72)

6.1 (6.1)

6.6 (6.6)

6.10 (6.10) Let $h \in C_H(S)$. Then, for all $s \in S$, $hs = sh$. That is for all $s \in S$, $h \in C_H(s)$, i.e. $h \in \bigcap_{s\in S} C_H(s)$. The converse is proven by following this logic in the other direction.

6.25 (6.25)

6.29 (6.29)

7.3 (7.3)

7.4 (7.4)

7.8 (7.8)

7.10 (7.10)

7.11 (7.11)

7.12 (7.12)

7.13 (7.13) Suppose σ is a homogenous permutation. Let $m \in \mathbb{Z}$. We need to prove that σ^m is a homogenous permutation. Of course the identity is homogenous, so we can assume that $\sigma^m \neq (1)$.

We now use previous exercise, Exercise 7.12. Let $N \in \mathbb{Z}$ be s.t. $(\sigma^m)^N \neq (1)$. Then since $(\sigma^m)^N = \sigma^{mN}$ is a power of σ, and since σ is assumed homogenous, we know that the support of $(\sigma^m)^N$ is equal to the support of σ. But σ^N is also a power of σ and so this common support also equals the support of σ^N and we are done.

7.17 (7.17) Two permutations have the same cycle type if (and only if) they are conjugate. But (in any group actually) the two elements

$$\sigma\tau \quad \text{and} \quad \tau\sigma$$

are conjugate (via σ^{-1}), indeed

$$\sigma^{-1}(\sigma\tau)\sigma \quad \text{and} \quad \tau\sigma.$$

7.18 (7.18) No solution can be given.

8.1 (8.1)

8.2 (8.2) This is actually hard to do wrong since any bijection which takes the identity to the identity will end up working. For instance the map

$$\begin{pmatrix} (0,0) & (1,0) & (0,1) & (1,1) \\ \mathbf{1} & r_1 & r_2 & \varrho_\pi \end{pmatrix}$$

which is a bijection. Replace every element of the table for $\mathbb{Z}_2 \times \mathbb{Z}_2$ by its image under the map gives us

$\mathbb{Z}_2^2$	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
$(0,0)$	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
$(1,0)$	$(1,0)$	$(0,0)$	$(1,1)$	$(0,1)$
$(0,1)$	$(0,1)$	$(1,1)$	$(0,0)$	$(1,0)$
$(1,1)$	$(1,1)$	$(0,1)$	$(1,0)$	$(0,0)$

$\rightsquigarrow$

$\mathbb{Z}_2^2$	$\mathbf{1}$	r_1	r_2	ϱ_π
$\mathbf{1}$	$\mathbf{1}$	r_1	r_2	ϱ_π
r_1	r_1	$\mathbf{1}$	ϱ_π	r_2
r_2	r_2	ϱ_π	$\mathbf{1}$	r_1
ϱ_π	ϱ_π	r_2	r_1	$\mathbf{1}$

.

This resulting table is literally the table for $S(▬)$ as it appears at the beginning of this chapter so it is a valid table and we have proved that this map gives us an isomorphism $\mathbb{Z}_2^2 \cong S(▬)$.

8.3 (8.3) The multiplication table for $\mathcal{D}_4$ is

$\mathcal{D}_4$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$
$\mathbf{1}$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$
ρ	ρ	ρ^2	ρ^3	$\mathbf{1}$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$
ρ^2	ρ^2	ρ^3	$\mathbf{1}$	ρ	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$
ρ^3	ρ^3	$\mathbf{1}$	ρ	ρ^2	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$
$\mathbf{r}$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{1}$	ρ	ρ^2	ρ^3
$\mathbf{r}\rho$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	ρ^3	$\mathbf{1}$	ρ	ρ^2
$\mathbf{r}\rho^2$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	ρ^2	ρ^3	$\mathbf{1}$	ρ
$\mathbf{r}\rho^3$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	ρ	ρ^2	ρ^3	$\mathbf{1}$

. If we call the map in question φ we have a new table

$\mathcal{D}_4$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}\rho$	$\mathbf{r}\rho\rho$	$\mathbf{r}\rho\rho^2$	$\mathbf{r}\rho\rho^3$
$\mathbf{1}$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}\rho$	$\mathbf{r}\rho\rho$	$\mathbf{r}\rho\rho^2$	$\mathbf{r}\rho\rho^3$
ρ	ρ	ρ^2	ρ^3	$\mathbf{1}$	$\mathbf{r}\rho\rho^3$	$\mathbf{r}\rho$	$\mathbf{r}\rho\rho$	$\mathbf{r}\rho\rho^2$
ρ^2	ρ^2	ρ^3	$\mathbf{1}$	ρ	$\mathbf{r}\rho\rho^2$	$\mathbf{r}\rho\rho^3$	$\mathbf{r}\rho$	$\mathbf{r}\rho\rho$
ρ^3	ρ^3	$\mathbf{1}$	ρ	ρ^2	$\mathbf{r}\rho\rho$	$\mathbf{r}\rho\rho^2$	$\mathbf{r}\rho\rho^3$	$\mathbf{r}\rho$
$\mathbf{r}\rho$	$\mathbf{r}\rho$	$\mathbf{r}\rho\rho$	$\mathbf{r}\rho\rho^2$	$\mathbf{r}\rho\rho^3$	$\mathbf{1}$	ρ	ρ^2	ρ^3
$\mathbf{r}\rho\rho$	$\mathbf{r}\rho\rho$	$\mathbf{r}\rho\rho^2$	$\mathbf{r}\rho\rho^3$	$\mathbf{r}\rho$	ρ^3	$\mathbf{1}$	ρ	ρ^2
$\mathbf{r}\rho\rho^2$	$\mathbf{r}\rho\rho^2$	$\mathbf{r}\rho\rho^3$	$\mathbf{r}\rho$	$\mathbf{r}\rho\rho$	ρ^2	ρ^3	$\mathbf{1}$	ρ
$\mathbf{r}\rho\rho^3$	$\mathbf{r}\rho\rho^3$	$\mathbf{r}\rho$	$\mathbf{r}\rho\rho$	$\mathbf{r}\rho\rho^2$	ρ	ρ^2	ρ^3	$\mathbf{1}$

=

$\mathcal{D}_4$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$
$\mathbf{1}$	$\mathbf{1}$	ρ	ρ^2	ρ^3	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$
ρ	ρ	ρ^2	ρ^3	$\mathbf{1}$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$
ρ^2	ρ^2	ρ^3	$\mathbf{1}$	ρ	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$
ρ^3	ρ^3	$\mathbf{1}$	ρ	ρ^2	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$
$\mathbf{r}\rho$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{1}$	ρ	ρ^2	ρ^3
$\mathbf{r}\rho^2$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	ρ^3	$\mathbf{1}$	ρ	ρ^2
$\mathbf{r}\rho^3$	$\mathbf{r}\rho^3$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	ρ^2	ρ^3	$\mathbf{1}$	ρ
$\mathbf{r}$	$\mathbf{r}$	$\mathbf{r}\rho$	$\mathbf{r}\rho^2$	$\mathbf{r}\rho^3$	ρ	ρ^2	ρ^3	$\mathbf{1}$

which can be quickly checked to be a valid table.

8.4 (8.4)

8.5 (8.5)

(1) Example 8.19**(7)** shows that in general, for any abelian group (which $\mathcal{U}(\mathbb{Z}_7)$ certainly is) and any integer n the map

$$\hat{}^n : a \mapsto a^n$$

is a group homomorphism. Here we are dealing with the case that $n = 3$.

(2) The kernel is

$$\begin{aligned}\operatorname{Ker}\left(\hat{}^3\right) &= \{a \in \mathcal{U}(\mathbb{Z}_7) : \hat{}^3(a) = 1\} \\ &= \{a \in \mathcal{U}(\mathbb{Z}_7) : a^3 = 1\}.\end{aligned}$$

There is no fancy trick to do this, but since the number involved are small we do this the old fashioned way (we use the representatives $-3, -2, -1, 1, 2, 3$) :

$$\begin{aligned} 1^3 = 1 \qquad & 2^3 = 8 \equiv_7 1 \qquad 3^3 = 27 \equiv_7 -1 \\ (-3)^3 = (-1)\cdot(3^3) \equiv_7 -(-1) = 1 \quad & (-2)^3 \equiv_7 -1 \qquad (-1)^3 = -1. \end{aligned} \tag{5}$$

Inspecting (5) shows us that $\operatorname{Ker}\left(\hat{}^3\right)$ is $\{[\,1\,]_7, [\,2\,]_7, \overset{[4]_7}{\overset{\shortparallel}{[\,-3\,]_7}}\}$.

(3) Again, inspecting (5) we now see that $\operatorname{Im}\left(\hat{}^3\right) = \{[\,1\,]_7, [\,-1\,]_7\}$.

(4) We have seen that the kernel has 3 elements and the image has 2. Moreover since $|\,\mathcal{U}(\mathbb{Z}_7)\,| = 6$ the statement is verified:

$$6 = 3 \cdot 2.$$

8.6 (8.6)

8.7 (8.7) Let $\varphi\colon G \to \mathbb{Z}_p$ be a homomorphism of groups. By the theorem right before this exercise, Theorem 8.12, we know that the image of φ, $\operatorname{Im}(\varphi)$, is a subgroup of $\mathbb{Z}_p$. But by Exercise/Theorem 6.26 we know that every subgroup of $\mathbb{Z}_p$, and hence $\operatorname{Im}(\varphi)$ is either all of $\mathbb{Z}_p$, in which case φ is surjective, or it (it being $\operatorname{Im}(\varphi)$) is trivial, in which case the map φ is trivial.

8.20 (8.20)

(1) Since the group $\mathcal{D}_4$ is not abelian, Exercise 8.12 back on page 384 says that the map $\hat{}^2$ is not a homomorphism.

(2) Let $w \in \mathcal{D}_4$ be arbitrary. We have $w^4 = \mathbf{1}_{\blacklozenge}$. Therefore $\hat{}^3(w) = w^3 = w^{-1}$. But again, Exercise 8.12 says that, as a non-abelian group, this is not a homomorphism.

(3) It's easy to check, using the fact that $\forall w \in \mathcal{D}_4, w^4 = \mathbf{1}_{\blacklozenge}$, that $\hat{}^5 = \mathbf{1}_{\mathcal{D}_4}$, which is of course a homomorphism.

8.23 (8.23) The homomorphism $\mathcal{A}_4 \to \mathbb{Z}_3$ discussed at the beginning of the chapter (detailed on page 375) is surjective, but since $\mathcal{A}_4$ is centerless and $\mathbb{Z}_3$ is abelian (hence its center is everything and contains all 3 elements of $\mathbb{Z}_3$) the center of $\mathcal{A}_4$ cannot possibly map onto the center of $\mathbb{Z}_3$.

8.30 (8.30) Since, by definition, the core is a subset of H the two parts are really the same. Now since $gHg^{-1} \leqslant G$ for every $g \in G$, the core is an intersection of subgroups of G and hence is a subgroup of G.

8.45 (8.45) The group $B\mathcal{D}_3$ contains an element of order 6, namely χ_6, but $\mathcal{A}_4$ does not. It only contains non-trivial elements of order 3 ($\underline{3}$–cycles) and 2 ($\underline{2}\,\underline{2}$–cycles). (Indeed, not even $\mathcal{S}_4$ contains an element of order 6.)

8.47 (8.47) If we do not want to base this on luck, note that $\hat{}^{10}\colon \mathcal{U}(\mathbb{Z}_{31}) \to \mathcal{U}(\mathbb{Z}_{31})$ is a homomorphism with image the set of elements of order 3. It's easy to see that $[\,2\,]_{31}$ is in the kernel of this map (which sucks), i.e. $[\,2^{10}\,]_{31} = [\,1\,]_{31}$ (actually even $2^5 = 32$ has this property), but we get lucky when we try the next largest element $[\,3\,]_{31}$

$$3^3 = 27 \equiv_{31} -4 \Rightarrow 3^6 = 16 \Rightarrow 3^9 \equiv_n -4 \cdot 16 = -2 \cdot 32 \equiv_{31} -2 \Rightarrow 3^{10} \equiv_{31} -6 = 25.$$

Since $25^2 \equiv_{31} 36 \equiv_{31} = 5$ this is our other element of order 3. Indeed, a direct check confirms that

$$5^3 = 125 = 4 \cdot 31 + 1 \Rightarrow 5^3 \equiv_{31} 1$$

and similarly for 25.

8.58 (8.58)

8.59 (8.59)Recall that the subgroup $\mathbb{Z}_m[n] \subseteq \mathbb{Z}_m$ which is defined as $\{[\,a\,]_m \in \mathbb{Z}_m : n\,[\,a\,]_m = [\,0\,]_m\}$ is cyclic (as are all subgroups of a cyclic group) of order $\gcd(n,m)$ and as such is isomorphism to $\mathbb{Z}_{\gcd(n,m)}$. Therefore we are done after we construct an isomorphism

$$\hom(\mathbb{Z}_n, \mathbb{Z}_m) \xrightarrow{\cong} \mathbb{Z}_m[n].$$

To this end, we first define a map

$$\begin{aligned}\Psi\colon\ \hom(\mathbb{Z}_n, \mathbb{Z}_m) &\to \mathbb{Z}_m\\ : \varphi &\mapsto \varphi([\,1\,]_n).\end{aligned}$$

$\operatorname{Im}(\Psi) = \mathbb{Z}_m[n]$ $\subseteq$ Since

$$n \cdot \varphi([\,1\,]_n) = \varphi(n\,[\,1\,]_n) = \varphi([\,n\,]_n) = \varphi([\,0\,]_n) = [\,0\,]_m$$

we have $\operatorname{Im}(\Psi) \subseteq \mathbb{Z}_m[n]$.

$\supseteq$ Let $[\,a\,]_m \in \mathbb{Z}_m[n]$ be arbitrary. We must construct a homomorphism φ so that $\Psi(\varphi_a) = [\,a\,]_m$. I claim that

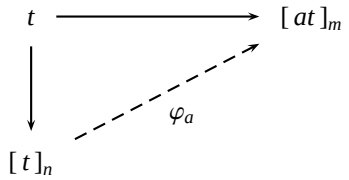
$$\varphi_a\colon [\,t\,]_n \mapsto [\,ta\,]_m$$

does the job. For starters, note that

$$\Psi(\varphi_a) = \varphi_a([\,1\,]_n) = [\,a\,]_m$$

so that much is done.

We still need to prove that the map φ_a is actually a homomorphism. Using Lemma 8.66 this isn't hard since we have that (for now anyway) the relation φ_a is induced as in the diagram



and the map upstairs, $t \mapsto [\,at\,]_m$ contains all of $n\mathbb{Z}$ in its kernel. This last fact is seen because for any $qn \in \mathbb{Z}_n\mathbb{Z}$ we have

$$[\,aqn\,]_m = q(n\,[\,a\,]_m) = q\,[\,0\,]_m = [\,0\,]_n$$

where we have used the assumption that $a \in \mathbb{Z}_m[n]$, i.e. $n \cdot [\,a\,]_m = [\,0\,]_m$. Since $\mathbb{Z} \to \mathbb{Z}_n$ has kernel equal to $\mathbb{Z}_n$, our (favorite!) lemma applies to prove that φ_a is a homomorphism. Thus completing the proof that $\operatorname{Im}(\Psi) \supseteq \mathbb{Z}_m[n]$.[IV]

Injective Using Example 8.72, we already know that both sides have the same order, namely

$$|\,\mathbb{Z}_m[n]\,| = \gcd(n,m) = |\,\hom(\mathbb{Z}_m, \mathbb{Z}_n)\,|$$

and we have just shown that Ψ surjective it is also injective (by the pigeon hole principle).

[IV] Careful! This can be confusing! We didn't yet show that the map Ψ is a homomorphism, what we did show was that for any $a \in \mathbb{Z}_m[n]$ the map defined by sending $[\,1\,]_n \mapsto [\,a\,]_m$ is a homomorphism!

OP [V] Let φ_1 and φ_2 be two group homomorphisms $\mathbb{Z}_n \to \mathbb{Z}_m$. Then

$$\begin{aligned}\Psi(\varphi_1 + \varphi_2) &= (\varphi_1 + \varphi_2)([\,1\,]_n) && \text{(Def'n of } \Psi\text{)}\\ &= \varphi_1([\,1\,]_n) + \varphi_2([\,1\,]_n) && \text{(Def'n of addition of functions)}\\ &= \Psi(\varphi_1) + \Psi(\varphi_2). && \text{(Def'n of } \Psi\text{)}\end{aligned}$$

Therefore Ψ is op completing the proof that it is an isomorphism, $\Psi\colon\ \hom(\mathbb{Z}_n, \mathbb{Z}_m) \to \mathbb{Z}_m[n]$.

9.15 (9.15)

9.16 (9.16)

9.17 (9.17)

9.19 (9.19) Let $\mathcal{L} := \operatorname{lcm}(|\mathcal{A}|)$. First note $\mathcal{L}$ as the lcm of the orders of the a_i's we have for each a_i, $|a_i|\,|\,\mathcal{L}$. In particular, $a_i^{\mathcal{L}} = e_A$. Now let $w \in A$ be arbitrary. We need to prove that $w^{\mathcal{L}} = e_A$. But as an element of A there exists integers $m_1 \ldots m_n$ and elements $a_1, \ldots, a_n$ of $\mathcal{A}$ so that

$$w = a_1^{m_1} \cdot \cdots \cdot a_n^{m_n}.$$

Therefore

$$\begin{aligned}w^{\mathcal{L}} &= \left(a_1^{m_1} \cdot \cdots \cdot a_n^{m_n}\right)^{\mathcal{L}} && (6)\\ &= (a_1^{m_1})^{\mathcal{L}} \cdot \cdots \cdot (a_n^{m_n})^{\mathcal{L}} && (A \text{ abelian implies } a \mapsto a^m \text{ is hom.})\\ &= (a_1^{\mathcal{L}})^{m_1} \cdot \cdots \cdot (a_n^{\mathcal{L}})^{m_n}\\ &= e_A^{m_1} \cdot \cdots \cdot e_A^{m_n}\\ &= e_A \cdot \cdots \cdot e_A\\ &= e_A.\end{aligned}$$

Remark L.1. Note that step is the one and only place that we used the fact that A is abelian.

9.20 (9.20)

9.21 (9.21) Since $X \subseteq \lfloor X \rceil$ (by **ngen** 1), and since $\lfloor X \rceil \trianglelefteq G$ (by **ngen** 2) ⟨in particular $\lfloor X \rceil \leqslant G$ ⟩, $\lfloor X \rceil$ satisfies **gen** 1 and **gen** 2, hence $\langle X \rangle \leqslant \lfloor X \rceil$ by **gen** 3.

9.22 (9.22) In $\mathcal{D}_4$, the subgroup $H := \{\mathbf{1}, \mathbf{r}\}$ (like any group) is normal in itself. Therefore

$$\lfloor H \rceil_H = H.$$

However, H is **not** a normal subgroup of $\mathcal{D}_4$. Therefore

$$\lfloor H \rceil_{\mathcal{D}_4} \neq H.$$

Indeed, it is actually equal $\left\langle \mathbf{r}, \mathbf{r}\rho^2 \right\rangle$ (which is a subgroup $\mathcal{D}_4$ whose order is 4).

[V] This may scare you, but let's take our time! It's not too bad if we stick with using our left brains! It helps to keep in mind the operation in both of the relevant groups, $\hom(\mathbb{Z}_n, \mathbb{Z}_m)$ and $\mathbb{Z}_n$ is addition.

9.23 (9.23)

9.24 (9.24)

ngen 1 Of course $g \in \mathrm{cc}(g)$.

ngen 2 Since for all $\widehat{g} \in G$, $\widehat{g}\mathrm{cc}g\widehat{g}^{-1} = \mathrm{cc}g$, $\langle\, \mathrm{cc}(g) \,\rangle \trianglelefteq G$.

ngen 3 If If $N \trianglelefteq G$ and $g \in N$, then for all $\widehat{g} \in G$, since $N \trianglelefteq G$ we still have that

$$\widehat{g}g\widehat{g}^{-1} \in N.$$

This means that

$$\mathrm{cc}(g) = \{\widehat{g}g\widehat{g}^{-1} : \widehat{g}\} \subseteq N.$$

As such, by **gen** 3 we have that $\langle\, \mathrm{cc}(g) \,\rangle \preccurlyeq N$.

9.26 (9.26) Let $g \in G$ be arbitrary. We need to prove that $gNg^{-1} = N$. But $gNg^{-1} = c_g(N)$. But since c_g is an automorphism and since N **char** G by assumption this is exactly what we have.

9.27 (9.27) Let $\varphi \in \mathrm{Aut}\,(G)$ be arbitrary. We need to prove that $\varphi(N_1) = N_1$. Since N_2 **char** G we know that $\varphi(N_2) = N_2$, and so $\varphi|_{N_2}$ is an automorphism of N_2, for emphasis let's call it ψ. But since N_1 **char** N_2 we know that $\psi(N_1) = N_1$. But of course $\psi(N_1) = \varphi(N_1)$ proving the result.

9.28 (9.27) Let $g \in G$ be arbitrary. Then conjugation by g, c_g, is an automorphism of G. Since $N_2 \trianglelefteq G$, c_g restricts to an automorphism of N_2. Since N_1 **char** N_2, we know that $c_g(N_1) = N_1$. This proves that for all $g \in G$, $gN_1g^{-1} = N_1$, i.e. $N_1 \trianglelefteq G$.

9.40 (9.40)

9.43 (9.43)

(1) If σ is even, then $F(\sigma) = \sigma$, which is still even. On the other hand, if σ is odd, then $F(\sigma) = \sigma \circ (n-1\ n)$ and, we still have

$$\varepsilon(F(\sigma)) = \varepsilon(\sigma \circ (n-1\ n)) = (-1)\cdot(-1) = 1.$$

So, in either case, $F(\sigma)$ is even and $\mathrm{Im}\,(F) \subseteq \mathcal{A}_n$.

(2) We need to show that F is operation preserving and that F is injective.

- First we show that F is op. To avoid using a bunch of cases, we use the formula

$$F(\sigma) = \sigma \circ (n-1\ n)^{\frac{1-\varepsilon(\sigma)}{2}}.$$

Note that $\sigma \mapsto (n-1\ n)^{\frac{1-\varepsilon(\sigma)}{2}}$ is a homomorphism which corresponds to the homomorphisms

$$\mathcal{S}_n \xrightarrow{\varepsilon} \{\pm 1\} \xrightarrow{\cong} \mathbb{Z}_2 \xrightarrow{\cong} \langle\, (n-1\ n) \,\rangle,$$

i.e. we have

$$(n-1\ n)^{\frac{1-\varepsilon(\sigma)}{2}}(n-1\ n)^{\frac{1-\varepsilon(\tau)}{2}} = (n-1\ n)^{\frac{1-\varepsilon(\sigma\tau)}{2}}. \tag{7}$$

For $\sigma, \tau \in \mathcal{S}_{n-2}$ we have

$$\begin{aligned} F(\sigma) \circ F(\tau) &= \sigma \circ (n-1\ n)^{\frac{1-\varepsilon(\sigma)}{2}} \circ \tau \circ (n-1\ n)^{\frac{1-\varepsilon(\tau)}{2}} \\ &= \sigma\tau \circ (n-1\ n)^{\frac{1-\varepsilon(\sigma)}{2}} (n-1\ n)^{\frac{1-\varepsilon(\tau)}{2}} && (\sigma \text{ and } \tau \text{ are disjoint from } ((n-1)\ n)) \\ &= \sigma\tau \circ (n-1\ n)^{\frac{1-\varepsilon(\sigma\tau)}{2}} && (\text{Equation } (7)) \\ &= F(\sigma\tau). \end{aligned}$$

- Suppose that $F(\sigma) = (1)$. Then $\sigma = (n-1\ n)^{\frac{1-\varepsilon(\sigma)}{2}}$. But since $\sigma \in \mathcal{S}_{n-2}$, the permutations σ and $(n-1\ n)$ are disjoint. So we must have that $\sigma = (1)$ proving that F is injective.

(3) We have a monomorphism $F : \mathcal{S}_n \to \mathcal{A}_{n+2}$, and like the image of any monomorphism $\mathcal{S}_n \cong F(\mathcal{S}_n)$.

10.10 (10.10)

10.21 (10.21)

10.22 (10.22)

10.25 (10.25)

10.26 (10.26)

(1)

(2) ~~⇒~~

⇐

(3) ~~⇒~~

⇐

(4)

10.27 (10.27)

(1)

(2)

(3)

(4) (No answer needed)

10.28 (10.28)

(1) For all $\mathbf{v}$ and $\mathbf{v}'$ in $\mathbb{F}^{2n}$ we have

$$\begin{aligned}
\widetilde{\mathbf{J}}(t^{\lambda}_{\mathbf{w}}(\mathbf{v}), t^{\lambda}_{\mathbf{w}}(\mathbf{v}')) &= \widetilde{\mathbf{J}}\left(\mathbf{v} + \lambda\widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w})\mathbf{w}, \mathbf{v}' + \lambda\widetilde{\mathbf{J}}(\mathbf{v}',\mathbf{w})\mathbf{w}\right) \\
&= \widetilde{\mathbf{J}}(\mathbf{v},\mathbf{v}') + \widetilde{\mathbf{J}}\left(\lambda\widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w})\mathbf{w}, \mathbf{v}'\right) + \widetilde{\mathbf{J}}\left(\mathbf{v}, \lambda\widetilde{\mathbf{J}}(\mathbf{v}',\mathbf{w})\mathbf{w}\right) + \widetilde{\mathbf{J}}\left(\lambda\widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w})\mathbf{w}, \lambda\widetilde{\mathbf{J}}(\mathbf{v}',\mathbf{w})\mathbf{w}\right) \\
&= \widetilde{\mathbf{J}}(\mathbf{v},\mathbf{v}') + \lambda\widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w}) \cdot \widetilde{\mathbf{J}}(\mathbf{w},\mathbf{v}') + \lambda\widetilde{\mathbf{J}}(\mathbf{v}',\mathbf{w}) \cdot \widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w}) + \lambda\widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w}) \cdot \lambda\widetilde{\mathbf{J}}(\mathbf{v}',\mathbf{w}) \cdot \cancel{\widetilde{\mathbf{J}}(\mathbf{w},\mathbf{w})} \\
&= \widetilde{\mathbf{J}}(\mathbf{v},\mathbf{v}') + \lambda\widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w}) \cdot \widetilde{\mathbf{J}}(\mathbf{w},\mathbf{v}') + \lambda\widetilde{\mathbf{J}}(\mathbf{v}',\mathbf{w}) \cdot \widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w}) \\
&= \widetilde{\mathbf{J}}(\mathbf{v},\mathbf{v}') + \lambda\widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w}) \cdot \widetilde{\mathbf{J}}(\mathbf{w},\mathbf{v}') - \lambda\widetilde{\mathbf{J}}(\mathbf{w},\mathbf{v}') \cdot \widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w}) \\
&= \widetilde{\mathbf{J}}(\mathbf{v},\mathbf{v}') + \cancel{\lambda\widetilde{\mathbf{J}}(\mathbf{w},\mathbf{v}') \cdot \widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w}) - \lambda\widetilde{\mathbf{J}}(\mathbf{w},\mathbf{v}') \cdot \widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w})} \\
&= \widetilde{\mathbf{J}}(\mathbf{v},\mathbf{v}') .
\end{aligned}$$

(2)

(3)

(4) We have that

$$\begin{aligned}
t^{\lambda^{-1}}_{\mathbf{w}-\mathbf{v}}(\mathbf{v}) &= \mathbf{v} + \lambda^{-1}\widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w}-\mathbf{v})(\mathbf{w}-\mathbf{v}) \\
&= \mathbf{v} + \lambda^{-1}\widetilde{\mathbf{J}}(\mathbf{v},\mathbf{w})(\mathbf{w}-\mathbf{v}) - \lambda^{-1}\cancel{\widetilde{\mathbf{J}}(\mathbf{v},\mathbf{v})}(\mathbf{w}-\mathbf{v}) \\
&= \mathbf{v} + \lambda^{-1}\lambda(\mathbf{v},\mathbf{w})(\mathbf{w}-\mathbf{v}) \\
&= \mathbf{v} + (\mathbf{w}-\mathbf{v}) \\
&= \mathbf{w}.
\end{aligned}$$

(5)

(6)

(7)

(8)

10.30 (10.30)

10.31 (10.31)

10.32 (10.32)

10.33 (10.33)

10.34 (10.34)

10.35 (10.35)

10.37 (10.37)

(1) Suppose that we had $kH = k'H$ for two elements k and $k' \in K$. This implies that

$$k^{-1}k' \in H.$$

But since $K \leqslant G$, we also have that $k^{-1}k' \in K$. Hence this product is in the intersection $H \cap K$. But this intersection is assumed to only contain the identity so that

$$k^{-1}k' = e_G \Rightarrow k' = k.$$

(2) In the first part we showed that

$$\{kH : k \in K\}$$

are all distinct cosets, therefore we've found $|K|$ cosets of G/H and so $|K| \leq |G/H|$. But by Lagrange's theorem,

$$|K| \leq \left|G/H\right| = \frac{|G|}{|H|} \Rightarrow |H| \cdot |K| \leq |G|.$$

The assumption that $|G| \leq |H| \cdot |K|$ therefore implies that all of the above inequalities are really equalities and so

$$|K| = \left|G/H\right|$$

as desired.

10.44 (10.44)

10.52 (10.52)

(1) Since there are 60 symmetries and 5 fix a pentagon, there are a total of $60/5 = 12$ pentagon faces in a soccer ball.

(2) Since there are 60 symmetries and 6 fix a hexagon, there are a total of $60/6 = 10$ hexagon faces in a soccer ball.

(3) Since there are 60 symmetries and only the identity (i.e. 1) fix a vertex, there are a total of $60/1 = 60$ vertices in a soccer ball.

(4) This is a joke question. Of course there are 0 square faces in a soccer ball.

10.68 (10.68)

10.69 (10.69) The number 31 is prime and the set of 3 by 3 matrices with determinant 1 and coefficients in $\mathbb{Z}_5$ is the group $\mathbf{SL}_3(\mathbb{Z}_5)$ which we has order 5616, (see Table 10.1)which is a number divisible by 31. Therefore Cauchy's lemma gives the result.

10.75 (10.75)

11.3 (11.3)

(1) Note that

$$x^2 = 1 \Leftrightarrow x^2 - 1 = 0 \Leftrightarrow (x-1) \cdot (x+1) = 0.$$

Since the ring $\mathbb{Z}_p$ is a field, this means that either $x - 1 = 0$ or $x + 1 = 0$, i.e. $x = 1$ or $x = -1$. Therefore $\mathbb{Z}_p^\times[2] = \{+1, -1\}$. So long as $p \neq 2$ these really are two different elements of $\mathbb{Z}_p$. Therefore, this part is a special case of the next.

(2) Since $\mathcal{U}(R)$ is a multiplicative abelian group, the map $\hat{}^2 \colon \mathcal{U}(R) \to \mathcal{U}(R)$ is a homomorphism. The kernel of this map is the set of $r \in \mathcal{U}(R)$ with $r^2 = 1$. Note that if $r \in R$ satisfies $r^2 = 1$, then r is invertible with (multiplicative) inverse equal to itself. Therefore

$$\operatorname{Ker}\left(\hat{}^2\right) = \{r \in \mathcal{U}(R) : r^2 = 1\} = \{r \in R : r^2 = 1\}.$$

Next note if $r \in R$ is an invertible perfect square, then there exists some $s \in R$ with $s^2 = r$. This implies that $s \cdot (sr^{-1}) = s^2 r^{-1} = 1$. In particular, s itself must also be invertible. Therefore the invertible perfect squares in R is the image of the map $\hat{}^2$. Putting this together using **KIT** gives us the desired result.

(3) We have

$$\mathbf{O}_2(\mathcal{U}(\mathbb{Z}_{7875})) = \mathbf{O}_2(\mathbb{Z}_8) \cdot \mathbf{O}_2(\mathbb{Z}_{124}) \cdot \mathbf{O}_2(\mathbb{Z}_6) = 2 \cdot 2 \cdot 2 = 8.$$

Therefore, there are a total of

$$\frac{8 \cdot 124 \cdot 6}{8} = 4 \cdot 62 \cdot 3 = 744$$

invertible perfect squares mod 7875.

(4) The number 5280 factors as $2^5 \cdot 3 \cdot 5 \cdot 11$. Therefore

$$\mathcal{U}(\mathbb{Z}_{5280}) \cong \mathcal{U}(\mathbb{Z}_{2^5}) \times \mathcal{U}(\mathbb{Z}_3) \times \mathcal{U}(\mathbb{Z}_5) \times U11 \cong \mathbb{Z}_2 \times \mathbb{Z}_8 \times \mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_{10}.$$

As such,

$$\mathbf{O}_2(5280) = \mathbf{O}_2(\mathbb{Z}_2 \times \mathbb{Z}_8 \times \mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_{10}) = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 32.$$

In conclusion, there are

$$1 \cdot 4 \cdot 1 \cdot 2 \cdot 5 = 40$$

invertible perfect squares in $\mathcal{U}(\mathbb{Z}_{5280})$.

(5) Of course 1, 4, $9 \equiv_{17} -8$, and $4^2 = 16 \equiv_{17} -1$ are perfect squares. Therefore the negative of every perfect square is a perfect square, as such $13 \equiv_{17} -4$, 8 are also perfect squares. Lastly, so is $8 \cdot 4 = 32 \equiv_{17} = -2$ and $-(-2) = 2$. In summary:

$$\mathbb{Z}_{17}^{\times 2} = \{1, 2, 4, 8, -8, -4, -2, -1\}.$$

11.7 (11.7)

11.10 (11.10) Define a map

$$\begin{aligned} \varphi &\colon G \times \mathbb{Z}_n \to G \\ &\colon (g, [\,m\,]_n) \mapsto gz^{-m}. \end{aligned}$$

The map φ is induced by the obvious map $\widehat{\varphi} \colon G \to \mathbb{Z}$, which is a homomorphism since

$$\begin{aligned} \widehat{\varphi}((g,m)(g',m')) &= \varphi((gg', m+m')) \\ &= gg'z^{-(m+m')}. \qquad = gg'z^{-m-m'} \end{aligned}$$

while at the same time

$$\begin{aligned} \widehat{\varphi}((g,m))\widehat{\varphi}((g',m')) &= gz^{-m}g'z^{-m'} \\ &= gg'z^{-m}z^{-m'} \quad (z \text{ is central}) \\ &= gg'z^{-m-m'}. \end{aligned}$$

Using the fact that our assumption is $|z| = n$, we see that φ is well-defined: $\widehat{\varphi}(e_G, n) = z^{-n} = e_G$.

The map is surjective: for any $g \in G$, $\varphi((g, [\,0\,]_n) = g$. Lastly, $(g, [\,m\,]_n) \in \mathrm{Ker}\,(\varphi)$ iff

$$\varphi((g, [\,m\,]_n)) = gz^{-m} = e_G \Leftrightarrow g = z^m \Rightarrow (g, [\,m\,]_n) = (z^m, [\,m\,]_n) = (z, [\,1\,]_n)^m.$$

Therefore $\mathrm{Ker}\,(\varphi) = \langle\, (z, [\,1\,]_n) \,\rangle$ and we can apply the first isomorphism theorem to conclude what we wanted.

11.25 (11.25)

(1) (By contradiction) If G/N were not perfect, there would be a non-trivial homomorphism of this quotient onto some abelian group A:

$$G/N \to A.$$

But composing this homomorphism with the quotient map, $G \to G/N$, gives us a non-trivial homomorphism $G \to A$ which is impossible since G is assumed perfect.

(2) Nope. For example, $B\left(\bigstar^3\right)$ is perfect and $\langle\, -1 \,\rangle$ is a normal subgroup, but $\langle\, -1 \,\rangle$ is not perfect (it's actually abelian).

11.26 (11.26) Let's denote the commutators $\{[x, x'] : x, x' \in S\}$ as $[S, S]$ and their normal closure as $\overline{\mathbf{S}}$. With this notation we need to prove that

$$[G, G] = \overline{\mathbf{S}}.$$

$\subseteq$ As a set of commutators we have that $[S, S] \leqslant [G, G]$. But since $[G, G] \trianglelefteq G$ we have that

$$\lfloor [S, S] \rceil = \overline{\mathbf{S}} \leqslant [G, G]\,.$$

$\supseteq$ To show the opposite inclusion, i.e.

$$[G, G] \leqslant \overline{\mathbf{S}}$$

we need to prove that the quotient $G_S := G/\overline{\mathbf{S}}$ is abelian. For starters, note that the quotient is a group since $\overline{\mathbf{S}}$ is normal in G. Next note that since G is assumed to generated S, the image of S under the quotient map

$$\begin{aligned} \pi_S : S &\to G_S \\ : s &\mapsto s\overline{\mathbf{S}} \end{aligned}$$

generates G_S. Therefore, to prove that G_S is abelian we simply need to prove that for every x and $y \in S$ the cosets $x\overline{\mathbf{S}}$ and $y\overline{\mathbf{S}}$ commute. Having established what we're really trying to prove, the actual proof is trivial! Indeed, by definition, $[x, y] \in [S, S] \subseteq \overline{\mathbf{S}}$.

11.29 (11.29)

(1) By the previous exercise, we know that

$$G^{perf} := \left\langle \bigcup H \right\rangle$$

(where the union is over all perfect subgroups of G) is perfect subgroup of G. This subgroup must contain all other perfect subgroups of G since if $K \leqslant G$ is perfect, then K must have been included in the union which defines G^{perf}.

(2) Let φ be an automorphism of G. Then $\varphi(G^{perf})$ is a perfect subgroup of G. Therefore by the maximal property of G^{perf}, this implies that

$$\varphi(G^{perf}) \subseteq G^{perf}.$$

Applying the same logic with the automorphism φ^{-1} replacing φ shows that

$$\varphi^{-1}(G^{perf}) \subseteq G^{perf}.$$

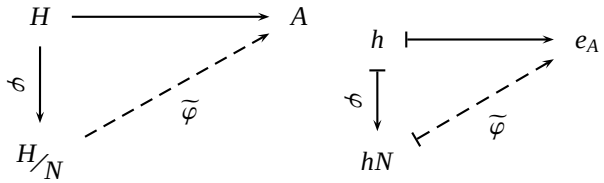
Taking φ of both sides of this equation shows us that $G^{perf} \subseteq \varphi(G^{perf})$ which implies that $G^{perf} = \varphi(G^{perf}$.

(3)

$\Rightarrow$ If G^{perf} is non-trivial, then G contains a non-trivial perfect subgroup. Since non-trivial perfect groups are not solvable, G contains a non- solvable subgroup and G itself cannot be solvable.

$\Leftarrow$ If G^{perf} is trivial, then G contains no non-trivial perfect subgroups. In particular G itself is not perfect, so $[G,G] \neq G$ and so G_{ab} is a non-trivial abelian group. Since any subgroup of $[G,G]$ is a subgroup of G, $[G,G]$ also does not contain a perfect subgroup. Since the order of $[G,G]$ is strictly less than the order of G we can proceed by induction to conclude that $[G,G]$ is solvable. Therefore we've found a solvable quotient with a solvable kernel and so Theorem 11.71**(2)** applies proving that G is also solvable.

(4) In general, suppose that N is a normal perfect subgroup of G and $H/N \leqslant G/N$ is a perfect subgroup of G/N for some subgroup H of G which contains N. I claim that H itself must be perfect. Indeed, if we have some homomorphism $\varphi\colon H \to A$ where A is abelian, then since N is assumed perfect, we must have that $\varphi|_N$ is trivial. This means that we have a well defined map induced map $H/N \to A$.



But this implies that φ is trivial. Therefore, in the special case that $N = G^{perf}$ we must have that H/G^{perf} is perfect implies that $H = G^{perf}$ and so the quotient is trivial. In particular, the quotient G/G^{perf} contains no proper perfect subgroups and so, by the previous part, this quotient must be solvable.

11.30 (11.30)

$\Leftarrow$ Since a non-trivial perfect group is not solvable, and since 11.71**(??)(??)** says that every subgroup of a solvable group is solvable we are done.

$\Rightarrow$ Now suppose that G is a finite non-solvable group. Then we must have an integer m so that the m^{th} terms in the derived series $G^{(m)}$ is non-trivial and $G^{(m)} = G^{(m+1)}$. By definition,

$$G^{(m+1)} = \left[G^{(m)}, G^{(m)}\right] \Rightarrow G^{(m)} = \left[G^{(m)}, G^{(m)}\right].$$

This proves that $G^{(m)}$ is a non-trivial perfect subgroup of G.

11.33 (11.33) If G were abelian, then $|(G/\mathcal{C}(G))| = 1$ which is not prime. OTOH, if G is non-abelian, then $G/\mathcal{C}(G)$ is not-cyclic. And since any group whose order is prime is cyclic, $|(G/\mathcal{C}(G))|$ cannot be prime in this case either.

11.34 (11.34) The center of $\mathcal{D}_4$ is ρ^2 and we've seen several times that $\mathcal{D}_4/\left\langle \rho^2 \right\rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$.

11.35 (11.35) Let G be a non-abelian group of order 16. Then $G/\mathcal{C}(G)$ can't have order 1 for then $\mathcal{C}(G)$ would equal G implying that G is abelian. If the quotient had order 2, then the quotient would be cyclic (isomorphic to $\mathbb{Z}_2$) which is also not allowed. Lastly the quotient can't have order 16 for in that case we'd have a 2–group with a trivial center.

11.42 (11.42)

11.43 (11.43)

11.44 (11.44)

11.45 (11.45)

11.46 (11.46) If G contained an index 2 subgroup, then we'd have a non-trivial homomorphism $G \to \mathbb{Z}_2$. But then if $\mathbb{Z}_2$ is generated by elements of odd order, then so would $\mathbb{Z}_2$. This is clearly false since every (well, the one) non-trivial element of $\mathbb{Z}_2$ has order 2, which is even.

13.1 (13.1)

$\Rightarrow$ If G is a p-group, meaning the order of G is p^n for some positive integer n, then by Lagrange's theorem every element of G has order which divides p^n. But the only divisors of p^n are powers of p all of which are divisible by p.

$\Leftarrow$ $\langle$ We prove this part by contraposition, i.e. we prove if G is not a p-group, then there exists an element whose order is not divisible by p.$\rangle$ Suppose G is not a p-group. Then there exists some prime q other than p which divides $|G|$. But by Cauchy's lemma G must contain an element of order q. Since p and q are distinct primes $p \nmid q$.

13.13 (13.13) The following table shows this (as usual, $\mathbf{o}_2(G)$ denotes the number of elements of order 2 in G)

G	$\lvert \mathcal{C}(G) \rvert$	$\mathbf{o}_2(G)$
$\mathbb{Z}_{12}$	12	1
$\mathbb{Z}_6 \times \mathbb{Z}_2$	12	3
$\mathcal{A}_4$	1	3
$\mathcal{D}_6$	2	7
$B\mathcal{D}_3$	2	1

because no two listed groups have both numbers the same. Since both columns are isomorphism invariants, none of the 5 listed groups can be isomorphic.

13.16 (13.16)

13.22 (13.22)

14.3 (14.3) Let G be a simple group of order 36. By **Sylow3**, we must have $n_3(G) = 4$. This gives us a non-trivial group homomorphism

$$G \to \mathcal{S}_4.$$

Since $|\mathcal{S}_4| = 24 < 36 = |G|$ this homomorphism cannot be injective and, as such, must have a non-trivial kernel, whence a PNT subgroup of G, a contradiction.

A.1 (A.1)

- I already gave away an example where the dp is not positive.
- Let R be a ring and n a positive integer. Then for any $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in R^n$ if $\mathbf{v} \neq \mathbf{0}$, then for some $i \in \overline{n}$ we must have $v_i \neq 0$. This means that $\mathbf{v} \bullet \mathbf{e}_i = v_i \neq 0$. As $\mathbf{v}$ was arbitrary, for ever vector there exists some vector with a non-zero dot product with the original, i.e. the dot product is non-degenerate as desired.

A.2 (A.2)

A.4 (A.4)

A.5 (A.5)

A.6 (A.6)

A.7

⇒ Suppose that we have $\mathcal{A}(\mathbf{V},\mathbf{V}) = \mathbf{0}$ for any vector $\mathbf{V} \in \mathbb{R}^3$ (that is suppose Property **(1)** holds). Now let $\mathbf{v},\mathbf{w} \in \mathbb{R}^3$ be arbitrary. Since $\mathbf{v}+\mathbf{w}$ is a vector in $\mathbb{R}^3$, we have that $\mathcal{A}(\mathbf{v}+\mathbf{w},\mathbf{v}+\mathbf{w}) = \mathbf{0}$ (this is Property **(1)** applied to $\mathbf{V} = \mathbf{v}+\mathbf{w}$). This implies that

$$\begin{aligned}
\mathbf{0} &= \mathcal{A}(\mathbf{v}+\mathbf{w},\mathbf{v}+\mathbf{w}) && (\textbf{(1)} \text{ applied to } \mathbf{V} = \mathbf{v}+\mathbf{w})\\
&= \mathcal{A}(\mathbf{v},\mathbf{v}) + \mathcal{A}(\mathbf{v},\mathbf{w}) + \mathcal{A}(\mathbf{w},\mathbf{v}) + \mathcal{A}(\mathbf{w},\mathbf{w}) && (\text{Bilinear property})\\
&= \mathbf{0} + \mathcal{A}(\mathbf{v},\mathbf{w}) + \mathcal{A}(\mathbf{w},\mathbf{v}) + \mathcal{A}(\mathbf{w},\mathbf{w}) && (\textbf{(1)} \text{ applied to } \mathbf{V} = \mathbf{v})\\
&= \mathcal{A}(\mathbf{v},\mathbf{w}) + \mathcal{A}(\mathbf{w},\mathbf{v}) + \mathbf{0} && (\textbf{(1)} \text{ applied to } \mathbf{V} = \mathbf{w})\\
&= \mathcal{A}(\mathbf{v},\mathbf{w}) + \mathcal{A}(\mathbf{w},\mathbf{v}).
\end{aligned}$$

Therefore, subtracting off,

$$\mathbf{0} = \mathcal{A}(\mathbf{v},\mathbf{w}) + \mathcal{A}(\mathbf{w},\mathbf{v}) \Rightarrow \mathbf{0} - \mathcal{A}(\mathbf{v},\mathbf{w}) = \mathcal{A}(\mathbf{w},\mathbf{v}) + \cancel{\mathcal{A}(\mathbf{v},\mathbf{w}) - \mathcal{A}(\mathbf{v},\mathbf{w})} \Rightarrow -\mathcal{A}(\mathbf{v},\mathbf{w}) = \mathcal{A}(\mathbf{w},\mathbf{v}).$$

Since $\mathbf{v}$ and $\mathbf{w}$ were arbitrary, we have proven that **(2)** holds whenever **(1)** does.

⇐ On the other hand assume **(2)** holds. We want to prove that $\mathcal{A}(\mathbf{V},\mathbf{V}) = \mathbf{0}$ for every $\mathbf{V} \in \mathbb{R}^3$.

$$\begin{aligned}
&\mathcal{A}(\mathbf{V},\mathbf{V}) = -\mathcal{A}(\mathbf{V},\mathbf{V}) && (\textbf{(2)} \text{ w/ } \mathbf{v} = \mathbf{V} \text{ and } \mathbf{w} = \mathbf{V})\\
\Rightarrow\ &\mathcal{A}(\mathbf{V},\mathbf{V}) + \mathcal{A}(\mathbf{V},\mathbf{V}) = -\mathcal{A}(\mathbf{V},\mathbf{V}) + \mathcal{A}(\mathbf{V},\mathbf{V})\\
\Rightarrow\ &2\mathcal{A}(\mathbf{V},\mathbf{V}) = \mathbf{0}\\
\Rightarrow\ &\mathcal{A}(\mathbf{V},\mathbf{V}) = \mathbf{0}.^{\text{VI}} && (\text{Cancel the 2})
\end{aligned}$$

Since $\mathbf{V}$ was arbitrary, we have proven that **(1)** holds whenever **(2)** does.

A.8 (A.8)

C.1 (C.1) No. The identity matrix $\mathbf{I}_n \in \mathbf{O}_n(\mathbb{R})$ but $2\mathbf{I}_n = \mathbf{I}_n + \mathbf{I}_n$ is never an orthogonal matrix since its first column is $2\mathbf{e}_1$ which have length 2, and not length 1 as is needed if $2\mathbf{I}_n$ did happen to be orthogonal.

C.2 (C.2) Let $\mathbf{A}$ be an orthogonal matrix.

⇒ Suppose in additional that $\mathbf{A}$ symmetric. Then $\mathbf{A}^\top = \mathbf{A}$. Moreover, since $\mathbf{A}$ is orthogonal, $\mathbf{A}^\top = \mathbf{A}^{-1}$. Combining these two facts shows that $\mathbf{A} = \mathbf{A}^{-1}$ or $\mathbf{A}^2 = \mathbf{I}$, as desired.

⇐ All implications can be reversed: if $\mathbf{A}^2 = \mathbf{I}$, then $\mathbf{A} = \mathbf{A}^{-1}$. And since $\mathbf{A}$ is orthogonal $\mathbf{A}^{-1} = \mathbf{A}^\top$ and so $\mathbf{A} = \mathbf{A}^\top$, i.e. $\mathbf{A}$ is symmetric.

C.5 (C.5)

[VI] This last step wouldn't work over $\mathbb{Z}_2$, but it would work for say $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{Z}_5$, $\mathbb{C}$, etc, etc.

(1) Suppose $\mathbf{v} \asymp \mathbf{v}'$. Then there exists a positive integer k so that $k\mathbf{v}' = \mathbf{v}$. For any $\mathbf{w} \in \mathbb{R}^n$ we therefore have

$$\pi_{\mathbf{v}}(\mathbf{w}) = \frac{\mathbf{v} \bullet \mathbf{w}}{\mathbf{v} \bullet \mathbf{v}}\mathbf{v} = \frac{k\mathbf{v}' \bullet \mathbf{w}}{k\mathbf{v}' \bullet k\mathbf{v}'}k\mathbf{v}' = \frac{k(\mathbf{v}' \bullet \mathbf{w})}{k^2(\mathbf{v}' \bullet \mathbf{v}')}k\mathbf{v}' = \frac{(\mathbf{v}' \bullet \mathbf{w})}{(\mathbf{v}' \bullet \mathbf{v}')}\mathbf{v}' = \pi_{\mathbf{v}'}(\mathbf{w}).$$

(2) This part follows at once from the fact that we have $\pi_{\mathbf{v}}(\mathbf{v}) = \mathbf{v}$.

(3) **(a)** Using the fact that $\mathbf{v} \bullet \mathbf{v} \neq 0$ we see that

$$\pi_{\mathbf{v}}(\mathbf{w}) = \frac{(\mathbf{v} \bullet \mathbf{w})}{(\mathbf{v} \bullet \mathbf{v})}\mathbf{v} \Rightarrow \pi_{\mathbf{v}}(\mathbf{w}) = \mathbf{0} \Leftrightarrow (\mathbf{v} \bullet \mathbf{w}) = 0 \Leftrightarrow \mathbf{v} \perp \mathbf{w}.$$

(b) If $\mathbf{w} = c\mathbf{v}$, then as already pointed out in the proof of 2, we have $\pi_{\mathbf{v}}(\mathbf{w}) = \mathbf{w}$. Conversely, if $\pi_{\mathbf{v}}(\mathbf{w}) = \mathbf{w}$, then using the formula

$$\mathbf{w} = \pi_{\mathbf{v}}(\mathbf{w}) = \frac{(\mathbf{v} \bullet \mathbf{w})}{(\mathbf{v} \bullet \mathbf{v})}\mathbf{v}$$

which implies that $\mathbf{v} \asymp \mathbf{w}$.

C.6 (C.6)

(1) We have that

$$\varrho_{\theta,\widehat{\mathbf{u}}} = \cos\theta\,\mathbf{v} + \sin\theta\,(\widehat{\mathbf{u}} \times \mathbf{v}).$$

Therefore

$$\begin{aligned}
\det(\mathbf{v}, \varrho_{\theta,\widehat{\mathbf{u}}}(\mathbf{v}), \widehat{\mathbf{u}}) &= \det(\mathbf{v}, \cos\theta\,\mathbf{v} + \sin\theta\,(\widehat{\mathbf{u}} \times \mathbf{v}), \widehat{\mathbf{u}}) \\
&= \cancel{\det(\mathbf{v}, \cos\theta\,\mathbf{v}, \widehat{\mathbf{u}})} + \det(\mathbf{v}, \sin\theta\,(\widehat{\mathbf{u}} \times \mathbf{v}), \widehat{\mathbf{u}}) && \text{(Dist and } \mathbf{v} \asymp \cos\theta\,\mathbf{v}) \\
&= \det(\mathbf{v}, \sin\theta\,(\widehat{\mathbf{u}} \times \mathbf{v}), \widehat{\mathbf{u}}) \\
&= \sin\theta \cdot \det(\mathbf{v}, \widehat{\mathbf{u}} \times \mathbf{v}, \widehat{\mathbf{u}}) && \text{(Scaler pull } \sin\theta) \\
&= \sin\theta \cdot \det(\widehat{\mathbf{u}}, \mathbf{v}, \widehat{\mathbf{u}} \times \mathbf{v}) && (\varepsilon(123) = 1) \\
&= \sin\theta \cdot (\widehat{\mathbf{u}} \times \mathbf{v}) \bullet (\widehat{\mathbf{u}} \times \mathbf{v}) && \text{(Def'n of CP)} \\
&= \sin\theta \cdot \left\| \widehat{\mathbf{u}} \times \mathbf{v} \right\|^2 \\
&= \sin\theta \cdot \left\| \widehat{\mathbf{u}} \right\|^2 \cdot \| \mathbf{v} \|^2 && (\widehat{\mathbf{u}} \perp \mathbf{v}) \\
&= \sin\theta \cdot \| \mathbf{v} \|^2 . && (\widehat{\mathbf{u}} \in \mathbb{S}^2)
\end{aligned}$$

(2) There are several ways of proving this:

Possibility 1 Since both $\mathbf{v}$ and $\varrho_{\theta,\widehat{\mathbf{u}}}(\mathbf{v})$ are perpendicular to $\widehat{\mathbf{u}}$, this cross product in question is parallel to $\widehat{\mathbf{u}}$. That is

$$\mathbf{v} \times \varrho_{\theta,\widehat{\mathbf{u}}}(\mathbf{v}) = \alpha\widehat{\mathbf{u}}.$$

By definition of the cross product,

$$\sin\theta \cdot \| \mathbf{v} \|^2 = \det(\mathbf{v}, \varrho_{\theta,\widehat{\mathbf{u}}}(\mathbf{v}), \widehat{\mathbf{u}}) = (\mathbf{v} \times \varrho_{\theta,\widehat{\mathbf{u}}}(\mathbf{v})) \bullet \widehat{\mathbf{u}} = \alpha\widehat{\mathbf{u}} \bullet \widehat{\mathbf{u}} = \alpha.$$

Possibility 2 Directly taking the cross product of $\mathbf{v}$ with $\varrho_{\theta,\widehat{\mathbf{u}}} = \cos\theta\mathbf{v} + \sin\theta\widehat{\mathbf{u}} \times \mathbf{v}$ we get

$$\mathbf{v} \times \varrho_{\theta,\widehat{\mathbf{u}}}(\mathbf{v}) = \sin\theta\,(\mathbf{v} \times (\widehat{\mathbf{u}} \times \mathbf{v}))\,.$$

Then we use the triple cross product formula (see Corollary 3.102 back on page 123) shows us that

$$\mathbf{v} \times (\widehat{\mathbf{u}} \times \mathbf{v}) = (\mathbf{v} \bullet \mathbf{v}) \cdot \widehat{\mathbf{u}} - \cancel{(\mathbf{v} \bullet \mathbf{u})} \cdot \mathbf{v} = \|\mathbf{v}\|^2 \cdot \widehat{\mathbf{u}}.$$

Plugging this in shows us (again!) that the formula indeed holds.

C.8 (C.8)

C.9 (C.9) We need to prove that $\mathbf{R}_\mathsf{q}$ preserves both scalar multiplication and addition.

(1) Let $\mathbf{v} \in \mathbb{R}^3$ and $c \in \mathbb{R}$. Then

$$\begin{aligned}\mathbf{R}_\mathsf{q}(c\mathbf{v}) &= \mathsf{q}c\mathbf{v} * \mathsf{q}^{-1} \\ &= c\mathsf{q}\mathbf{v} * \mathsf{q}^{-1} \\ &= c\mathbf{R}_\mathsf{q}(\mathbf{v}).\end{aligned}$$

Therefore, $\mathbf{R}_\mathsf{q}$ preserves scalar multiplication.

(2) Let $\mathbf{v}$ and $\mathbf{w} \in \mathbb{R}^3$, then

$$\begin{aligned}\mathbf{R}_\mathsf{q}(\mathbf{v}+\mathbf{w}) &= \mathsf{q}(\mathbf{v}+\mathbf{w})\mathsf{q}^{-1} \\ &= (\mathsf{q}(\mathbf{v}+\mathbf{w}))\mathsf{q}^{-1} \\ &= (\mathsf{q}\mathbf{v}+\mathsf{q}\mathbf{w})\mathsf{q}^{-1} \\ &= \mathsf{q}\mathbf{v}\mathsf{q}^{-1} + \mathsf{q}\mathbf{w}\mathsf{q}^{-1} \\ &= \mathbf{R}_\mathsf{q}(\mathbf{v}) + \mathbf{R}_\mathsf{q}(\mathbf{w}).\end{aligned}$$

Therefore $\mathbf{R}_\mathsf{q}$ preserves addition.

C.14 (C.14) This is a straight forward "plug and chug" using the formula

$$\begin{pmatrix} a^2+b^2-c^2-d^2 & 2bc-2ad & 2ac+2bd \\ 2ad+2bc & a^2-b^2+c^2-d^2 & 2cd-2ab \\ 2bd-2ac & 2ab+2cd & a^2-b^2-c^2+d^2 \end{pmatrix}$$

Bringing out the common $\frac{1}{6}^2$ term:

$$\frac{1}{6^2}\begin{pmatrix} 1^2+5^2-3^2-1^2 & 2\cdot5\cdot3-2\cdot1\cdot1 & 2\cdot1\cdot3+2\cdot5\cdot1 \\ 2\cdot1\cdot1+2\cdot5\cdot3 & 1^2-5^2+3^2-1^2 & 2\cdot3\cdot1-2\cdot1\cdot5 \\ 2\cdot5\cdot1-2\cdot1\cdot3 & 2\cdot1\cdot5+2\cdot3\cdot1 & 1^2-5^2-3^2+1^2 \end{pmatrix} = \frac{1}{36}\begin{pmatrix} 16 & 28 & 16 \\ 32 & -16 & -4 \\ 4 & 16 & -32 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 4 & 7 & 4 \\ 8 & -4 & -1 \\ 1 & 4 & -8 \end{pmatrix}.$$

C.15 (C.15)

C.18 (C.18) We hit the vector $\begin{bmatrix}1\\1\\1\end{bmatrix}$ around until we get the 20 distinct vertices of . The powers of $\mathbf{R}_{q_5}$ behave as follows on this vector:

$$\mathbf{R}_{q_5}\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right)=\begin{bmatrix}\phi\\0\\\phi^{-1}\end{bmatrix} \qquad \mathbf{R}_{q_5}\left(\begin{bmatrix}\phi\\0\\\phi^{-1}\end{bmatrix}\right)=\begin{bmatrix}\phi\\0\\-\phi^{-1}\end{bmatrix}$$

$$\mathbf{R}_{q_5}\left(\begin{bmatrix}\phi\\0\\-\phi^{-1}\end{bmatrix}\right)=\begin{bmatrix}1\\1\\-1\end{bmatrix} \qquad \mathbf{R}_{q_5}\left(\begin{bmatrix}1\\1\\-1\end{bmatrix}\right)=\begin{bmatrix}\phi^{-1}\\\phi\\0\end{bmatrix}.$$

Of course a much easier calculation shows $\mathbf{R_i}$ takes the original vector $\begin{bmatrix}1\\1\\1\end{bmatrix}$ to $\begin{bmatrix}1\\-1\\-1\end{bmatrix}$ (of course you should be able to do that much by sight!). Now let's use $\mathbf{R}_{q_5}$ to hit around this resulting vector:

$$\mathbf{R}_{q_5}\left(\begin{bmatrix}1\\-1\\-1\end{bmatrix}\right)=\begin{bmatrix}0\\\phi^{-1}\\-\phi\end{bmatrix} \qquad \mathbf{R}_{q_5}\left(\begin{bmatrix}0\\\phi^{-1}\\-\phi\end{bmatrix}\right)=\begin{bmatrix}-\phi^{-1}\\\phi^{-1}\\0\end{bmatrix}$$

$$\mathbf{R}_{q_5}\left(\begin{bmatrix}-\phi^{-1}\\\phi^{-1}\\0\end{bmatrix}\right)=\begin{bmatrix}0\\\phi^{-1}\\\phi\end{bmatrix} \qquad \mathbf{R}_{q_5}\left(\begin{bmatrix}0\\\phi^{-1}\\\phi\end{bmatrix}\right)=\begin{bmatrix}1\\-1\\1\end{bmatrix}.$$

Like the icosahedron, the dodecahedron is has the anti-polar map as an (orientation reversing) symmetry. We could fiddle with this longer, but we can instead simply note that applying the antipolar map to the ten vertices we have found so far gives us 10 new vertices giving us a total of 20 vertices. In summary, the 20 vertices of the dodecahedron are

$$\mathbf{V}_1=\begin{bmatrix}1\\1\\1\end{bmatrix} \quad \mathbf{V}_2=\begin{bmatrix}\phi\\0\\\phi^{-1}\end{bmatrix} \quad \mathbf{V}_3=\begin{bmatrix}\phi\\0\\-\phi^{-1}\end{bmatrix} \quad \mathbf{V}_4=\begin{bmatrix}1\\1\\-1\end{bmatrix} \quad \mathbf{V}_5=\begin{bmatrix}\phi^{-1}\\\phi\\0\end{bmatrix}$$

$$\mathbf{V}_6=\begin{bmatrix}1\\-1\\-1\end{bmatrix} \quad \mathbf{V}_7=\begin{bmatrix}0\\\phi^{-1}\\-\phi\end{bmatrix} \quad \mathbf{V}_8=\begin{bmatrix}-\phi^{-1}\\\phi^{-1}\\0\end{bmatrix} \quad \mathbf{V}_9=\begin{bmatrix}0\\\phi^{-1}\\\phi\end{bmatrix} \quad \mathbf{V}_{10}=\begin{bmatrix}1\\-1\\1\end{bmatrix}$$

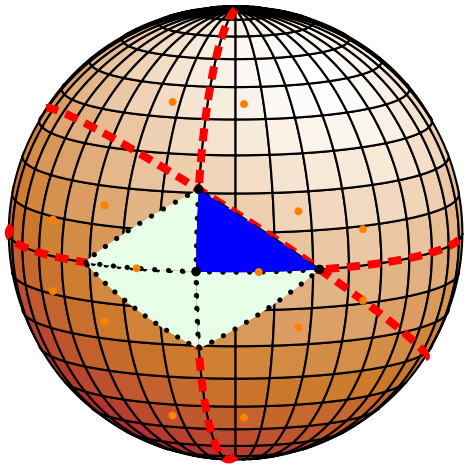
$$\mathbf{V}_{11}=\begin{bmatrix}-1\\-1\\-1\end{bmatrix} \quad \mathbf{V}_{12}=\begin{bmatrix}-\phi\\0\\-\phi^{-1}\end{bmatrix} \quad \mathbf{V}_{13}=\begin{bmatrix}-\phi\\0\\\phi^{-1}\end{bmatrix} \quad \mathbf{V}_{14}=\begin{bmatrix}-1\\-1\\1\end{bmatrix} \quad \mathbf{V}_{15}=\begin{bmatrix}-\phi^{-1}\\-\phi\\0\end{bmatrix}$$

$$\mathbf{V}_{16}=\begin{bmatrix}-1\\1\\1\end{bmatrix} \quad \mathbf{V}_{17}=\begin{bmatrix}0\\-\phi^{-1}\\\phi\end{bmatrix} \quad \mathbf{V}_{18}=\begin{bmatrix}\phi^{-1}\\-\phi^{-1}\\0\end{bmatrix} \quad \mathbf{V}_{19}=\begin{bmatrix}0\\-\phi^{-1}\\-\phi\end{bmatrix} \quad \mathbf{V}_{20}=\begin{bmatrix}-1\\1\\-1\end{bmatrix}.$$

C.19 (C.19)

E.8 (E.8) The first three parts are a special case of the {where}. For the last part consider the usual ring homomorphism

$$\mathbb{Z}\to\mathbb{Z}_3$$



{To Do}

taking every integer to it's class mod 3. This homomorphism of rings induces

$$\mathbb{Z}_2 \cong \mathcal{U}(\mathbb{Z}) \to \mathcal{U}(\mathbb{Z}_3) \cong \mathbb{Z}_2$$

and isomorphism of groups but it is clearly, but it clearly is not an isomorphism of rings since $\mathbb{Z}$ is infinite and $\mathbb{Z}_3$ is not.

E.9 (**E.9**) Let $n \in \mathbb{Z}$ and $\varphi\colon \mathbb{Z} \to \mathbb{Z}$ be a ring homomorphism. Then since $\varphi(1) = 1$ by definition of a ring homomorphism,

$$\varphi(n) = \varphi(1 + \cdots + 1) = n\varphi(1) = n.$$

Since n was arbitrary, φ is the identity function.

E.13 (**E.13**) The statement that $\mathbb{Z}_3 \preccurlyeq \mathbb{Z}_9$ is false. Indeed, it is tempting to say that

$$\mathbb{Z}_3 = \{[\,0\,]_9\,, [\,0\,]_9\,, [\,2\,]_9\},$$

but this is certainly false. For instance $[\,1\,]_9 + [\,2\,]_9]$ is not zero. (The multiplicative part is correct, but the additive part is wrong.) Another common mistake is to say that

$$\mathbb{Z}_3 = \{[\,0\,]_9\,, [\,3\,]_9\,, [\,6\,]_9\},$$

Now the additive part is matches up correctly, but the multiplicative part is wrong. Actually, we're going to show that the additive part of any field of order 9 is must be isomorphic to $\mathbb{Z}_3 \times \mathbb{Z}_3$.

J.2 (**J.2**) Note that $|\,\sigma\,| = 4$, $|\,\tau_1\,| = |\,\tau_2\,| = 2$, σ is central and

$$\tau_2\tau_1\tau_2 = \sigma^2\tau_1.$$

Therefore G, like G_{13} is a semi-direct product $\mathbb{Z}_4 \times \mathbb{Z}_2$ by $\mathbb{Z}_2$ and so $G_{13} \cong G$. (We never needed to directly prove that G has order 16)

J.4 (**J.4**)

J.5 (**J.5**)

J.6 (J.6)

J.7 (J.7)

J.8 (J.8)

J.9 (J.9) The answer to this question is G_8. If I wanted to be a jerk, I could simply write out the 16 elements of G_8 and tell you to verify that it's a group and leave you to find the order table. However, since I rock and all, I will explain to you the though process that I used.

Let P_2 denote the (a) Sylow 2–subgroup of $\mathbf{GL}_2(\mathbb{Z}_3)$. In Example 10.76, we saw that $\mathbf{SL}_2(\mathbb{Z}_3)$ is isomorphic to $B\left(\blacktriangleright^3\right)$. As such, $\mathbf{SL}_2(\mathbb{Z}_3)$, and hence also $\mathbf{GL}_2(\mathbb{Z}_3)$, contains (an isomorphic copy of) $B\left(\blacktriangleright^3\right)$ which in turn contains (an isomorphic copy of) $3\mathrm{l}Q_8$. Hence P_2 contains (an isomorphic copy of) Q_8, call it $\dot{Q}_8$. Note that since $Q_8 \trianglelefteq B\left(\blacktriangleright^3\right)$ we have that $\dot{Q}_8 \trianglelefteq \mathbf{SL}_2(\mathbb{Z}_2)$ and so $\dot{Q}_8$ is the one and only copy of Q_8 in $\mathbf{SL}_2(\mathbb{Z}_2)$, and any pair of two order 4 matrices generate $\dot{Q}_8$. We take the matrices

$$\dot{\mathbf{i}} := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad \dot{\mathbf{j}} := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

for no other reason other than the fact that these were the first two order four matrices I could think of! (Don't forget! Any traceless matrix with determinant 1 and trace 0 has order 4). Note that in accordance to the isomorphism, $Q_8 \cong \dot{Q}_8$, the matrix

$$\dot{\mathbf{i}} \cdot \dot{\mathbf{j}} = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

has order 4. Naturally enough we denote this matrix $\dot{\mathbf{k}}$.

Next note that we have an additional element of order 2 (besides $-\mathbf{I}_2$) in P_2 since, since for instance, the permutation matrix

$$\mathbf{F} := \mathbf{F}_{(1\,2)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

is a matrix which has order 2 and no conjugate of it is not in $\dot{Q}_8$ since this latter group, as a subgroup of $\mathbf{SL}_2(\mathbb{Z}_3)$, consists of matrices having determinant 1. A direct calculation shows that $\mathbf{F}$ normalizes $\dot{Q}_8$.[VII] Indeed,

$$\mathbf{F} \cdot \dot{\mathbf{i}} \cdot \mathbf{F}^{-1} = -\dot{\mathbf{i}} \qquad \mathbf{F} \cdot \dot{\mathbf{j}} \cdot \mathbf{F}^{-1} = \dot{\mathbf{k}} \qquad \mathbf{F} \cdot \dot{\mathbf{k}} \cdot \mathbf{F}^{-1} = \dot{\mathbf{j}}.$$

As such,

$$\left\langle \dot{\mathbf{i}}, \dot{\mathbf{j}}, \mathbf{F} \right\rangle = \left\langle \dot{\mathbf{i}}, \dot{\mathbf{j}} \right\rangle \langle \mathbf{F} \rangle \Rightarrow \left| \left\langle \dot{\mathbf{i}}, \dot{\mathbf{j}}, \mathbf{F} \right\rangle \right| = 16.$$

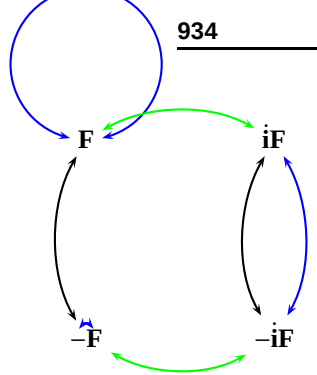
Therefore we have

$$P_2 = \left\langle \dot{\mathbf{i}}, \dot{\mathbf{j}}, \mathbf{F} \right\rangle.$$

Being a copy of Q_8, which we are quite familiar with, we already know what the order of each of the 8 elements of $\dot{Q}_8$ are. We have the 8 elements of $P_2 \setminus \dot{Q}_8$ left to consider, the most obvious such element is the involution $\mathbf{F}$. Let's check out the the conjugacy class of $\mathbf{F}$.

[VII] Again, from general theory we knew this would work even before we even try it. This is due to the fact that

$$n_2(\mathbf{SL}_2(\mathbb{Z}_2)) = 1 \Rightarrow \dot{Q}_8 \textbf{ char } \mathbf{SL}_2(\mathbb{Z}_3) \Rightarrow \dot{Q}_8 \trianglelefteq \mathbf{GL}_2(\mathbb{Z}_3) \Rightarrow \mathbf{F} \text{ normalizes } \dot{Q}_8.$$



where

- **blue** denotes conjugation by **F**
- **black** denotes conjugation by $\dot{\mathbf{i}}$
- **green** denotes conjugation by $\dot{\mathbf{j}}$.

(So, counting $-\dot{1}$ and these 4 elements, we have $\mathbf{o}_2(P_2) \geq 5$). We have 4 elements left to consider. Whose left? Well, the element $\dot{\mathbf{j}}\mathbf{F}$ satisfies

$$\begin{aligned}(\dot{\mathbf{j}}\mathbf{F})^2 &= \dot{\mathbf{j}} \cdot \mathbf{F} \cdot \dot{\mathbf{j}} \cdot \mathbf{F} \\ &= \dot{\mathbf{j}} \cdot \left(\mathbf{F} \cdot \dot{\mathbf{j}} \cdot \mathbf{F}^{-1}\right) \\ &= \dot{\mathbf{j}} \cdot \dot{\mathbf{k}} \\ &= \dot{\mathbf{i}}.\end{aligned}$$

Therefore, since $\dot{\mathbf{i}}$ has order 4, $\dot{\mathbf{j}}\mathbf{F}$ has order 8. And since 3, 5, and 7 are relatively prime to 8 those respective powers

$$(\dot{\mathbf{j}}\mathbf{F})^3, \qquad (\dot{\mathbf{j}}\mathbf{F})^5, \quad \text{and} \quad (\dot{\mathbf{j}}\mathbf{F})^7$$

of $\dot{\mathbf{j}}\mathbf{F}$ also have order 8. This shows that the remaining four elements of P_2 have order 8. Therefore, P_2 has the order table

$\mathbf{P}_2$	n	1	2	4	8	16
	$\mathbf{o}_n$	1	5	6	4	0

Inspecting the order tables presented, the only one which matches is the one for G_8 and in conclusion,

$$G_8 \cong P_2.$$

K.1 (**K.1**) Using the binomial theorem, we see that

$$(a + bp^i)^{p^j} = a^p + p^j \cdot a^{p-1}(bp^j) + C \cdot p^{i+j} = a^p + a^{p-1}b \cdot p^{i+j} + C \cdot p^{i+j} \equiv_{p^{i+j}} a^p$$

as promised.

Bibliography

[DF04] David S. Dummit and Richard M. Foote. Abstract algebra. John Wiley & Sons Inc., Hoboken, NJ, third edition, 2004. 1, E.62

[Gar08] Paul B. Garrett. Abstract algebra. Chapman & Hall/CRC, Boca Raton, FL, 2008. 1

[Hun80] Thomas W. Hungerford. Algebra, volume 73. Springer-Verlag, New York, 1980. Reprint of the 1974 original. 1

[Rot96] Joseph J. Rotman. A first course in abstract algebra. Prentice Hall Inc., Upper Saddle River, NJ, 1996. 1

[SSA01] Maurice; Smith, Douglas; Eggen and Richard St. Andre. Transition to Advanced Mathematics. Brokes/Cole, Pacific Grove, CA, 2001. 1, 2.3, IX, 7

Index

$\|\mathbf{v}\|$, 743
$\prod$ notation., 8
n-sphere, 746
$\binom{n}{m}$, 12
Inn (G), 460

Abelian, 230
Action
  Definition, 479
  Faithful, 482
  Fixed Point Set, 482, 483
  Free, 482
  Kernel, 482
  n-Transitive, 494
  Noun, 479
  Orbit, 482
  Stabilizer
    Of a Point, 482
    Of a Set, 483
  Transitive, 482, 494
  Verb, 479
Algebraic Integers, 820
Algebraic System, 213
  Sub-algebraic system, 220
Alternating Group, 41
Associativity, 218
Aut (G), 381
Automorphism
  Inner, 460

Bijection, 32
Bilinear, 745
Binary Groups
  Binary n-gon group, 304
  Binary Icosahedral, 387
  Binary Octahedral, 387
  Binary Tetrahedral, 387
Binary Operation, 213

Cartesian Product, 8
Cayley Action, 486
Cayley Hamilton Theorem, 201
Cayley Hamilton Theorem, 2 by 2, 201
Center
  $\mathcal{S}_n$ centerless, 356
  $\mathcal{D}_n$ centerless for n-odd, 356
  Definition, 326
  When Quotient is Cyclic, 642
Centralizier
  Definition, 326
Characteristic Subgroup, 459
Closure, 220
Commute, 230
Complement, 6
Composition of Relations, 20
Conjugate
  Definition, 297
  Equivalence Relation, 298
Conjugation Trick, 360
Cosets, 520
  Left, 520
  Right, 530
Cut and Paste, 360
Cycle Decomposition of a Permutation, 344
Cycles, 344
  Disjoint, 356
Cyclic Group, 279

Definition of a Group, 237
Derived Length, 638
Determinant, 180
Diagonal Matrix, 210

Dictionary, 39, 401
Digraph, 18
Dihedral Group, 245
Domain, 24

Equivalence Relation, 48
    Equivalence Class, 54
    Generated by a Relation, 52
    Give Partitions, 57
    Quotient, 54
    Quotient Function, 55
Euler Totient, 153

Fiber Product, 613
Field, 152, 814
Fractional Linear Transformation, 850
Function, 28
    Bijection, 32
    Corestriction, 29
    Image, 24
    Image and Inverse Image, 44
    Injective, 25
    Invertible, 32
    Restriction, 29
    Section, 28
Functions
    Definition, 28
    Equal, 31
Fundamental Theorem of Algebra, 839
$\mathcal{F}_X$, 35

$\langle g \rangle$, 279
General Linear Group, 205
Generalized Associativity
    Of Relations, 22
Golden Ratio, 100
Greatest Common Divisor, 76
Group
    Definition, 228, 237
Group Action
    Quotient, 498

Heisenberg group, 274
Hermitian Conjugate, 323
hom(G, A), 382
Homomorphism
    Automorphism, 372
    Definition, 370
    Endomorphism, 372
    Epimorphism, 372
    First Examples, 382
    Isomorphism, 370
    Monomorphism, 372
Homomorphism of Rings, 821

icositetrachoron, 573
Ideal
    Maximal, 828
    Prime, 828
Image, 24
Index, 524
Infinite Dihedral Group, 547
Inner Automorphisms, 460
Integers mod n, 56
Integral Domain, 814
Intersection, 6
Irreducible, 815
Isomorphism, 370

Ker (φ), 377
Kernel, 377
Klein 4 group, 254

Lagrange's Theorem, 531
    Converse, 531
    Corollaries, 531
Landau Function., 365
Least Common Multiple, 76
Left Cosets, 520
Length of a vector, 743
Loop, 225

Magma, 213
Malnormal Subgroup, 405
Matrix
    Addition, 168
    Anti-Symmetric, 197
    Definition, 166
    Diagonal, 210
    Identity, 169
    Identity Matrix, 169
    Multiplication

Definition, 168
Is Associative, 175
Is Distributive, 170
Scalar, 210
Symmetric, 197
Trace, 199
Transpose, 193
Maximal ideal, 828
Monoid, 224
Multiplication, 213

n-Transitive Subgroups, 495
Normal Subgroup, 405
Normalizer Subgroup, 411

Orbit Stabilizer Theorem, 553
Order, 241
Ordered Pair, 8
OST, 553

p-group
Definition, 581, 681
Parallel, 50
Partial Order, 48
Partition
Dual, 706
Pauli Spin Matrices, 876
Perfect Group, 633
Perfect Square, 610
Permutation, 35
Braid Diagram, 41
Conjugation Trick, 360
Cut and Paste, 360
Cycle Type, 350
Even, 41
Fixed Point Set, 343
Odd, 41
Parity
Redefinition, 398
Sign, 41
Support, 343
Permutation Matrices, 389
Permutations
Recycling, 347
φ, 100
Prime ideal, 828
Principle Ideal, 827
Principle Ideal Domain, 827
Projective Linear Group, 240, 617

Q_8, 306
Quaternion
$\mathbb{R}^3$, 105
$\mathbb{S}^1$, 114
$\mathbb{S}^2$, 114
$\mathbb{S}^3$, 114
Conjugate, 112
Definition, 104
Pure, 105
Pure Part, 105
Q_8, 306
Real, 105
Real Part, 105
Section, 104
Unit, 114

Reducible, 815
Relation
Defined Everywhere, 25
Definition, 16
Domain, 24
Equivalence Relation, 48
First Set, 16
Image, 24
Injective, 25
Partial Order, 48
Reflexive, 48
Second Set, 16
Single-Valued, 25
Symmetric, 48
Transitive, 48
Relations
Equal, 31
Relatively Prime Integers, 77
Right Action, 486
Right Coset, 530
Ring
Subring, 811
Roots of Unity, 98
Rubik's Cube, 439

Scalar Matrix, 210

Self Normalizing, 412
Semi-Group, 218
Simple Group
    Definition, 409
Skew Field, 814
Skew–Field, 152
Solvability Chain, 638
Solvable, 638
Special Linear Group, 205
Standard Basis Vector, 742
Subgroup
    Definition, 311
    Malnormal, 403
    Normal, 403
    Tests, 313
Subgroups, 312
    Subgroup Test, 314
Subset, 4
Superset, 4
$\mathcal{S}_X$, 35
Sylow subgroup
    Definition, 682
Symplectic Group, 512
Symplectic Matrix, 512

Trace, 199
Transpose, 193
Trivial Kernel, 380

Union, 6
Unitary Group, 323

V, 254
Vierergruppe, 254

Well Ordering Principle, 78
Well-Defined, 64
Wilson's Theorem, 310

Young Diagram, 706

Zero Vector, 741