Math 453: Group Theory Fact Sheet

March 4, 2019

Contents

0 A few things 4

1 Very brief review of set theory 6

2 Binary operations on sets 8

2.1 Basic concepts 8

2.2 Algebraic Structures 8

2.3 Compatible Functions for Algebraic Structures 9

3 Groups: General concepts 10

3.1 Notation and equations manipulation in groups 10

3.2 Homomorphisms 11

3.3 Properties of elements 12

3.4 Subgroups 12

3.5 Centralizers and the center of a group 13

3.6 Conjugation of subgroups 14

3.7 Subgroups generated by subsets and cyclic subgroups 15

3.8 Subgroup series 16

3.8.1 The commutator subgroup 16

3.8.2 Lower central series 16

3.8.3 Derived series 16

3.9 Some classes of groups 16

3.9.1 Nilpotent groups 16

3.9.2 Solvable groups 16

3.9.3 *p*–groups 17

4 Product sets, quotient groups, and the isomorphism theorems 18

4.1 Inverse and product sets in groups . . . 18
4.1.1 Inverse and products sets . . . 18
4.1.2 Monoid structure on the power set $\mathscr{P}$ of G . . . 18
4.1.3 Products with subgroups . . . 19
4.1.4 Other identities . . . 19
4.2 Cosets of a subgroup . . . 19
4.2.1 Concrete example: Lines in $\mathbf{R}^2$ . . . 19
4.2.2 Cosets . . . 25
4.2.3 Normality of subgroups via coset conditions . . . 25
4.3 Coset and quotient spaces . . . 25
4.3.1 Equivalence relations via subgroups . . . 25
4.3.2 Equivalence classes associated subgroups . . . 26
4.3.3 Quotient sets and quotient maps . . . 27
4.3.4 Binary operations on cosets spaces . . . 27
4.3.5 Normal subgroups and quotient groups . . . 29
4.3.6 Summary of quotients . . . 30
4.3.7 Index of a subgroup . . . 30
4.4 The Isomorphism Theorems . . . 31
4.4.1 The First Isomorphism Theorem . . . 31
4.4.2 The Second Isomorphism Theorem . . . 34
4.4.3 The Third Isomorphism Theorem . . . 35
5 Groups: Examples 37
5.1 Basics of $\mathbf{R}^n$ and linear functions . . . 37
5.2 Basics: Matrices and Linear Functions . . . 37
5.2.1 Vector addition and scalar multiplication for $\mathrm{M}(n,m;\mathbf{R})$. . . . 39
5.2.2 Vector addition and scalar multiplication for $\mathrm{Hom}(\mathbf{R}^m,\mathbf{R}^n)$. . . . 39
5.2.3 Dot product or Euclidean inner product . . . 39
5.2.4 Matrix muliplication . . . 39
5.2.5 Square matrices . . . 40
5.3 The general and special linear groups . . . 40
5.3.1 Special types of matrices . . . 41
5.3.2 $\mathrm{O}(2)$ and $\mathrm{SO}(2)$ . . . 42
5.3.3 A geometric view of orthogonal matrices . . . 42
5.3.4 The translations group: $\mathrm{Trans}(\mathbf{R}^2)$ . . . 44

5.3.5 The affine group: $\text{Aff}(\mathbf{R}^2)$ . . . 44
5.3.6 Some explicit subgroups of $\text{GL}(2,\mathbf{R})$: group of upper triangular matrices . . . 46
5.3.7 Some subgroups of upper triangular matrices . . . 48
5.3.8 Higher dimensions: Heisenberg group . . . 50
5.3.9 Higher dimensions: Orthogonal and special orthogonal groups . . . 51
5.4 $\text{Sym}(m)$: The symmetric group on m objects . . . 51
5.4.1 Special elements . . . 52
5.4.2 Decompositions of elements . . . 52
5.4.3 Conjugacy classes, normalizers, and centralizers . . . 52
5.4.4 The alternating group: $\text{Alt}(m)$ . . . 52
5.4.5 Other special subgroups . . . 52
5.4.6 One realization of the symmetric group by matrices . . . 52
5.5 Free groups . . . 52
6 Group Actions 54
6.1 Basics . . . 54
6.2 The Orbit-Stabilizer Theorem . . . 56

0 A few things

This note contains definitions and basic facts from the theory of groups. It can serve as a first reference or bare bones starting point. Text that appears in blue has a hyperlink (typically to wikipedia).

Axiom of choice: In all that follows, I will assume the axiom of choice and will use it (mostly) without saying it. A thousand apologies for that.

Suggestions for the reader:

The layout of these notes is such that the note should not necessarily be read linearly. The first part of the notes (beyond the set theory review) deal with basic definitions and concepts for groups. The second part deals more with explicit examples and constructions of groups. The third part of the notes cover the concept of group actions. The fourth part on groups will focus on finite groups and covers some standard results from this topic. As such, the reader will likely want to move around more so as the basic definitions will be something the new reader will likely refer to with great frequency.

The hyperlinks throughout this note should be opened and at least fairly well skimmed. Some of the hyperlinks will likely be opened several times while others will have less utility. I view the links and their information as an important component of this note.

As with any note on any sufficiently complex topic, the reader should consider consulting regularly several other books/notes. J. S. Milne has several class notes on his webpage. Keith Conrad also has several useful notes on his webpage; one good collection is here. Below are some links to class notes in abstract algebra and foundational mathematics that I found in a quick search:

(1) Notes on abstract algebra. These notes look like a solid introduction to basic algebra.

(2) Abstract algebra. These notes cover a lot of different topics (well beyond what we can do in one term).

(3) Abstract algebra. This notes are very basic and short. That is not a bad thing when you are starting out.

(4) An introduction to algebra. These notes are perhaps a bit on the unusual or quirky side. The material covered looked nice and there is a section on methods of proof.

(5) Basic concepts in math. These notes cover basics in logic, proof methods, .etc.

(6) Abstract algebra: Theory and applications. The notes cover all of the basic material and more. These notes also talk about applications in cryptography and has SAGE programming exercises. These notes might be great for people that like coding, though they look like very good notes for any student starting out.

(7) Groups, rings, fields. These short notes cover a good bit of material, though they seem a bit more faster paced than some of the others.

(8) Abstract algebra. This relative short class notes covers most/all of the material we will cover in our class. There are some applications in cryptography and a few other excursions (e.g. wallpaper designs) in the note that makes it look really good.

(9) The art of proof. These notes cover all of the basic concepts that serve as the foundations of mathematics. Method of proof, axioms, logic, induction/recursion, and a great deal more are covered.

Remark: The above notes were found using google. I only looked at around 25-30 webpages (if that). There are lots of other notes and references out there. For instance, there is a wikipedia style webpage on properties and concepts in groups called Groupprops. This webpage is very handy for finding definitions in group theory. I wish such things had existed when I was a student.

Remark: I obviously have not read completely the notes linked above. In the event that one of the above notes says something wrong (in whatever sense), please know that I am not responsible for the content in the above links. If you should read something that you find troubling (e.g. personally or mathematical), let me know immediately. I also presume that accessing the above material for personal use is legal. If this is not the case, please let me know.

1 Very brief review of set theory

Given a set X, the cardinality of X is denoted by $|X|$ which loosely is the number of elements in the set. Finite sets have cardinality that is some natural number n. Given two sets, $\mathrm{Fun}(X,Y)$ is the set of all functions from X to Y. When $X=Y$, we denote the set of function in this case simply by $\mathrm{Fun}(X)$. The set of injective function $f\colon X \to Y$ will be denoted by $\mathrm{Inj}(X,Y)$ and when $X=Y$ by $\mathrm{Inj}(X)$. The set of all surjective functions $f\colon X \to Y$ will be denoted by $\mathrm{Sur}(X,Y)$ and when $X=Y$ by $\mathrm{Sur}(X)$. Finally, we denote the set of all bijective functions $f\colon X \to Y$ by $\mathrm{Bi}(X,Y)$ and when $X=Y$ by $\mathrm{Bi}(X)$. When $|X|=n$, we will also denote $\mathrm{Bi}(X)$ in this case by $\mathrm{Sym}(n)$.

Remark: Provided $X,Y \neq \emptyset$, the set $\mathrm{Fun}(X,Y) \neq \emptyset$. However, the subsets $\mathrm{Inj}(X,Y)$ or $\mathrm{Sur}(X,Y)$ could be empty. For instance, if $X=\{x\}$ and $Y=\{y_1,y_2\}$, then $\mathrm{Sur}(X,Y)=\emptyset$ and $\mathrm{Inj}(X,Y)=\mathrm{Fun}(X,Y)$. Since $\mathrm{Bi}(X,Y) \subset \mathrm{Inj}(X,Y)$ and $\mathrm{Bi}(X,Y) \subset \mathrm{Sur}(X,Y)$ holds for any sets X,Y, the set $\mathrm{Bi}(X,Y)$ can also be empty.

Remark: One brief aside on cardinalities, we say sets X,Y have the same cardinality when $\mathrm{Bi}(X,Y) \neq \emptyset$. In this case, we write $|X|=|Y|$. The following is a theorem of Schröder–Bernstein theorem:

Theorem Let X,Y be sets. The following are equivalent:

(a) $\mathrm{Bi}(X,Y) \neq \emptyset$.

(b) $\mathrm{Inj}(X,Y) \neq \emptyset$ and $\mathrm{Inj}(Y,X) \neq \emptyset$.

The implication (a) → (b) is easy. The implication (b) → (a) is the hard part. In fact, we see that all of the following statements are equivalent (these equivalences require the axiom of choice):

- $\mathrm{Bi}(X,Y) \neq \emptyset$.
- $\mathrm{Inj}(X,Y) \neq \emptyset$ and $\mathrm{Inj}(Y,X) \neq \emptyset$.
- $\mathrm{Sur}(X,Y) \neq \emptyset$ and $\mathrm{Sur}(Y,X) \neq \emptyset$.
- $\mathrm{Inj}(X,Y) \neq \emptyset$ and $\mathrm{Sur}(X,Y) \neq \emptyset$.

If $\mathrm{Inj}(X,Y) \neq \emptyset$, then we write $|X| \leq |Y|$. If $\mathrm{Sur}(X,Y) \neq \emptyset$, then we write $|X| \geq |Y|$. In this notation, we see that the above four set statements can be rewritten in this notation. The following four statements are equivalent:

- $|X|=|Y|$.
- $|X| \leq |Y|$ and $|Y| \leq |X|$.
- $|X| \geq |Y|$ and $|Y| \geq |X|$.
- $|X| \leq |Y|$ and $|X| \geq |Y|$.

To end this discussion with something concrete, take $X=\{a,b\}$, $Y=\{\alpha,\beta,\gamma\}$, and $Z=\{1,2,3,4\}$. Then we see that

$$\mathrm{Inj}(X,Y),\mathrm{Inj}(X,Z),\mathrm{Inj}(Y,Z) \neq \emptyset$$

and

$$\mathrm{Sur}(Z,X),\mathrm{Sur}(Z,Y),\mathrm{Sur}(Y,X) \neq$$

and

$$\mathrm{Bi}(X,Y)=\mathrm{Bi}(X,Z)=\mathrm{Bi}(Y,Z)=\emptyset.$$

Hence, we see that

$$|X| < |Y| < |Z|.$$

Of course, $|X| = 2$, $|Y| = 3$, and $|Z| = 4$. So we proved that $2 < 3 < 4$.

It is clear that if $|X|, |Y| < \infty$, then $|\mathrm{Fun}(X,Y)|, |\mathrm{Fun}(Y,X)| < \infty$. If $|X| = m$ and $|Y| = n$, then the size of the sets

$$\mathrm{Fun}(X,Y), \mathrm{Fun}(Y,X), \mathrm{Inj}(X,Y), \mathrm{Inj}(Y,X), \mathrm{Sur}(X,Y), \mathrm{Sur}(Y,X), \mathrm{Bi}(X,Y), \mathrm{Bi}(X,Y)$$

depend only on the integers m,n. Here are some fun counting problems:

Let $|X| = m$ and $|Y| = n$. We will assume that $m \neq n$. Determine the sizes of the followings sets as a function of m,n.

- $\mathrm{Fun}(X,Y)$ and $\mathrm{Fun}(Y,X)$.
- $\mathrm{Inj}(X,Y)$ and $\mathrm{Inj}(Y,X)$.
- $\mathrm{Sur}(X,Y)$ and $\mathrm{Sur}(Y,X)$.
- $\mathrm{Bi}(X,Y)$ and $\mathrm{Bi}(Y,X)$.

Some of the above are rather easy to count. Some of the above are less easy to count (e.g. use the inclusion-exclusion principle). I appreciate some of you have more experience than others. For those with a bit less experience, let me do one of the easy counts. Let us count the size of the set $\mathrm{Fun}(X,Y)$ when $|X| = m$ and $|Y| = n$. The formula does not depend on the relationship between m,n (i.e. whether $m < n$, $m = n$, or $m > n$). A function from X to Y assigns to each $x \in X$, some $y \in Y$. If it is easier for you, think of X as $\{1,2,\dots,m\}$ and $Y = \{1,2,\dots,n\}$. Regardless how you view X and Y, to build a function from X to Y, for each $x \in X$, we must choose some $y \in Y$. For each $x \in X$, the number of choices we have is $|Y| = n$. Since we are not imposing any additional conditions on our function (e.g. injective or surjective), this means that the number of functions from X to Y is $|Y|^{|X|}$ or n^m. We can check this for small values of m,n easily. For instance if $m = 1$ and $n = 2$, then we have two (2^1) functions

$$\begin{aligned} 1 &\mapsto 1 \\ 1 &\mapsto 2. \end{aligned}$$

When $m = 2$ and $n = 2$, we have four (2^2) functions

$$\begin{aligned} &1 \mapsto 1, \quad 2 \mapsto 1 \\ &1 \mapsto 1, \quad 2 \mapsto 2 \\ &1 \mapsto 2, \quad 2 \mapsto 1 \\ &1 \mapsto 2, \quad 2 \mapsto 2. \end{aligned}$$

When $m = 3$ and $n = 2$, we have eight (2^3) functions

$$\begin{aligned} &1 \mapsto 1, \quad 2 \mapsto 1, \quad 3 \mapsto 1 \\ &1 \mapsto 1, \quad 2 \mapsto 1, \quad 3 \mapsto 2 \\ &1 \mapsto 1, \quad 2 \mapsto 2, \quad 3 \mapsto 1 \\ &1 \mapsto 1, \quad 2 \mapsto 2, \quad 3 \mapsto 2 \\ &1 \mapsto 2, \quad 2 \mapsto 1, \quad 3 \mapsto 1 \\ &1 \mapsto 2, \quad 2 \mapsto 1, \quad 3 \mapsto 2 \\ &1 \mapsto 2, \quad 2 \mapsto 2, \quad 3 \mapsto 1 \\ &1 \mapsto 2, \quad 2 \mapsto 2, \quad 3 \mapsto 2. \end{aligned}$$

2 Binary operations on sets

Given a set X, we will be interested in functions from the set $X \times X$ to the set X where $X \times X$ denotes the Cartesian product. Usually when we work with functions, we label these $f \colon X \times X \to X$. However, since we want to view these functions as a type of operation on the set X that is analogous to adding or multiplying real numbers, we will use a notation similar to what we use in that setting. With this said, we now list some of the basic definitions in group theory.

2.1 Basic concepts

Definition 1 (Binary operation)**.** A **binary operation** on a set X is a function $\star \colon X \times X \to X$.

Definition 2 (Associative operator)**.** We say that a binary operation $\star$ on a set X is **associative** if for all $x, y, z \in X$, we have

$$z \star (y \star x) = (z \star y) \star x.$$

Definition 3 (Left/Right identity for an operator)**.** Given a set X and a binary operator $\star$ on X, we say that $x_0 \in X$ is a **left identity** for $\star$ if for all $x \in X$, we have $x_0 \star x = x$. We say that $x_0 \in X$ is a **right identity** for $\star$ if for all $x \in X$, we have $x \star x_0 = x$.

Definition 4 (Identity for an operator)**.** Given a set X and a binary operator $\star$ on X, we say that $x_0 \in X$ is an **identity** for $\star$ if x_0 is both a left and right identity for $\star$.

Remark: Given a set X with a binary operator $\star$, if $\star$ has an identity then this identity is unique (this is on Pset 1). Consequently, we will denote the identity for $\star$ by $1_{X,\star}$ or simple 1_X.

Definition 5 (Left/Right inverses)**.** Let X be a set with a binary operator $\star$ and an identity 1_X for $\star$. We say that $y \in X$ is a **left inverse** of $x \in X$ if $y \star x = 1_X$. We say that $y \in X$ is a **right inverse** of $x \in X$ if $x \star y = 1_X$. We say that $y \in X$ is a **inverse** of $x \in X$ if y is both a left and right inverse for x

Remark: If $x \in X$ has an inverse y for $\star$, then y is unique. Consequently, we usually denote y by x^{-1}.

Definition 6 (Inverses for binary operators)**.** Let X be a set with a binary operator $\star$ and identity element 1_X for $\star$. We say that $\star$ has **left inverses** if each $x \in X$ has a left inverse for $\star$. We say $\star$ has **right inverses** if each $x \in X$ has a right inverse for $\star$. We say that $\star$ has inverses if each $x \in X$ has an inverse for $\star$.

Definition 7 (Commutative operator)**.** Given a set X and a binary operator $\star$ on X, we say that $\star$ is **commutative** if for all $x, y \in X$, we have $x \star y = y \star x$.

Definition 8 (Products of operators)**.** If X is a set with a binary operator $\star$ and Y is a set with a binary operator $*$, the **product operator** is defined to be a binary operator $\star_\times * \colon (X \times Y) \times (X \times Y) \to X \times Y$ on the product set $X \times Y$ defined by

$$(x_1, y_1) \star_\times * (x_2, y_2) = (x_1 \star x_2, y_1 * y_2).$$

2.2 Algebraic Structures

Definition 9 (Magma)**.** A **magma** is a pair $(M, \star)$ where M is a set and $\star$ is a binary operator on M.

Definition 10 (Semigroup)**.** A **semigroup** is a pair $(S, \star)$ where S is a set and $\star$ is an associative binary operator on S.

Definition 11 (Unital Magma)**.** A **unital magma** is a triple $(M, \star, 1_M)$ where M is a set, $\star$ is a binary operator on M, and $1_M \in M$ is an identity for $\star$.

Definition 12 (Monoid)**.** A **monoid** is a triple $(M,\star,1_M)$ where M is a set, $\star$ is an associative binary operator on M, and $1_M \in M$ is an identity for $\star$.

Definition 13 (Group)**.** A **group** is a triple $(G,\star,1_G)$ where G is a set, $\star$ is an associative binary operator on G, 1_G is an identity for $\star$, and $\star$ has inverses.

Definition 14 (Commutative Group)**.** We say a group $(G,\star,1_G)$ is **commutative** if $\star$ is a commutative operator.

Remark: In honor of the mathematician Abel, commutative groups are also called **abelian groups**. Indeed, it is much more common to refer to the group as an abelian group than as a commutative group (see the wiki link above for commutative as evidence).

Remark: The product of any of the above objects is done with the binary operator being the product of binary operators given above. The identity in the product is given by taking the identity in each of the sets in the product as the 2–tuple.

2.3 Compatible Functions for Algebraic Structures

Definition 15 (Homomorphism: Magmas)**.** Let $(M,\star)$ and $(N,*)$ be magmas. We say that a function $f\colon M \to N$ is a **homomorphism** (or **magma homomorphism**) if for all $m_1,m_2 \in M$, we have

$$f(m_1 \star m_2) = f(m_1) * f(m_2).$$

Definition 16 (Homomorphisms: Unital Magmas)**.** Let $(M,\star,1_M)$ and $(N,*,1_N)$ be unital magmas. We say that a function $f\colon M \to N$ is a **homomorphism of unital magmas** if f is a magma homomorphism and $f(1_M) = 1_N$.

Remark: The definition for a homomorphism when the algebraic structures are semigroups is the same as that for magmas.

Remark: The definition for a homomorphism when the algebraic structures are monoids is the same as that for unital magmas.

Definition 17 (Homomorphism: Groups)**.** Let $(G,\star,1_G)$ and $(H,*,1_H)$ be groups. We say that a function $\varphi\colon G \to H$ is a **homomorphism** (or group homomorphism) if for all $g_1,g_2 \in G$, we have

$$\varphi(g_1 \star g_2) = \varphi(g_1) * \varphi(g_2).$$

Remark: If $\varphi\colon G \to H$ is a homomorphism of groups, then $\varphi(1_G) = 1_H$. You will prove this on a problem set.

3 Groups: General concepts

In the above sections, we have used a lot of notation to denote a group G. In group theory, one does not typically use this much notation in part because of laziness.

3.1 Notation and equations manipulation in groups

- Notation for the binary operator: Given a group $(G, \star, 1_G)$ and $g, h \in G$, we will write
$$gh = g \star h.$$
Because we will not explicitly use the $\star$ to denote the group operation, we can reduce our notation for a group to be a pair $(G, 1_G)$.

- Notation for inverses and powers: Given a group $(G, 1_G)$ and $g \in G$, we write g^{-1} for the inverse of g and
$$g^n = \underbrace{gg \dots g}_{n \text{ times}}.$$
Hence, we have
$$g^2 = gg, \quad g^5 = ggggg, \quad g^{-3} = g^{-1}g^{-1}g^{-1}.$$
We also have that
$$g^m g^n = g^{m+n}.$$
Hence, we have
$$g^2 g^3 = g^5, \quad g^{-7} g^9 = g^2, \quad g^{101} g^{-100} = g.$$

- Notation for identity: Given a group $(G, 1_G)$, we typically denote the identity element simply by 1. When more than one group is involved, for instance, when $\varphi \colon G \to H$ is a homomorphism, we would write $\varphi(1) = 1$ with the understanding that the 1 on the left side is really 1_G and the 1 on the right side is really 1_H. In particular, we will now simply refer to G as a group without mentioning the identity element 1_G and the binary operator $\star$.

- Commutator: Given a group G and $g, h \in G$, the **commutator** of g, h is
$$[g, h] = g^{-1}h^{-1}gh \in G.$$

- General notation: Given a group $g, h \in G$, we can make other elements of G by taking combinations of g, h. For instance, we can take the element
$$g^2 h^{-3} g h^5 g^{-4} h^{11}$$
or
$$ghghghghghghghghghghghghghghghghghgh.$$
It could be that $g^3 = 1$ or $gh^7 = g^2$. With this type of view, we could ask equational questions. For instance, given $g, h \in G$, is there an $x \in G$ such that
$$xg^3 = gh^3 x?$$
We could take elements $\gamma, \eta, \theta \in \Gamma$, where Γ is a group. We can build elements like
$$\gamma^2 \eta^5 \theta^{-2} \gamma \theta \gamma^{-8} \eta^{101}.$$

If you do not like the notation γ, η, θ, we can simply declare $x = \gamma$, $y = \eta$, and $z = \theta$. It is all just symbols that can be changed and manipulated. You just have to remember that

$$ghgh \neq g^2h^2$$

in most groups. Indeed, notice that

$$ghgh = (gh)^2.$$

In particular, we have

$$(gh)^2 \neq g^2h^2$$

in most groups.

3.2 Homomorphisms

We will now restate the definition of a group homomorphism in our new, simplified notation.

Definition 18 (Group Homomorphism)**.** Given groups G, H, we say $\varphi\colon G \to H$ is a **homomorphism** if

$$\varphi(g_1g_2) = \varphi(g_1)\varphi(g_2)$$

holds for all $g_1, g_2 \in G$.

Let $\mathbf{R}^+$ denote the positive real numbers. The set $\mathbf{R}^+$ can be viewed as a group under multiplication.

For each $\alpha \in \mathbf{R}^+$, the classical logarithm function $\log_\alpha$ base α or the natural logarithm function ln are functions $\mathbf{R}^+ \to \mathbf{R}$ satisfying

$$\log_\alpha(xy) = \log_\alpha(x) + \log_\alpha(y), \quad \ln(xy) = \ln(x) + \ln(x).$$

In particular, viewing $\mathbf{R}$ as a group under addition, $\log_\alpha, \ln$ are group homomorphisms.

For each $\alpha > 0$, we have the exponential function α^x and for $\alpha = e$, the function $\exp\colon \mathbf{R} \to \mathbf{R}^+$ where $\exp(x) = e^x$. The functions $\alpha^x, \exp$ satisfies

$$\alpha^{x+y} = \alpha^x\alpha^y, \quad \exp(x+y) = \exp(x)\exp(y).$$

In particular, $\alpha^x, \exp$ are group homomorphisms.

Definition 19 (Kernel)**.** Given groups G, H and a homomorphism $\varphi\colon G \to H$, the **kernel** of φ is defined to be

$$\ker\varphi = \{g \in G \ : \ \varphi(g) = 1\}.$$

Definition 20 (Image)**.** Given groups G, H and a homomorphism $\varphi\colon G \to H$, the **image** of φ is defined to be

$$\varphi(G) = \{h \in H \ : \ \varphi(g) = h \text{ for some } g \in G\}.$$

Definition 21 (Graph of a homomorphism)**.** Given groups G, H and a homomorphism $\varphi\colon G \to H$, the **graph of φ** is the subgroup of $G \times H$ defined by

$$\mathrm{Graph}(\varphi) = \{(g, \varphi(g)) \in G \times H \ : \ g \in G\}.$$

Definition 22 (Trivial homomorphism)**.** Given groups G, H, the **trivial homomorphism** of G to H is the homomorphism $\varphi_{\mathrm{triv}}\colon G \to H$ defined by $\varphi_{\mathrm{triv}}(g) = 1$ for all $g \in G$.

Definition 23 (Isomorphism)**.** Given groups G, H and a homomorphism $\varphi\colon G \to H$, we say φ is an **isomorphism** if φ is bijective.

Definition 24 (Isomorphic)**.** Given groups G,H, we say that G and H are **isomorphic** if there exists an isomorphism $\varphi\colon G \to H$. When G,H are isomorphic, we will denote this by $G \cong H$.

The $\exp, \ln$ are isomorphisms, and so $\mathbf{R}^+$ under multiplication and $\mathbf{R}$ under addition are isomorphic.

Definition 25 (Automorphism)**.** Given a group G, we will call an isomorphism $\varphi\colon G \to G$ an **automorphism**.

Given a pair of groups, we define $\mathrm{Hom}(G,H)$ to be the set of all homomorphisms $\varphi\colon G \to H$. Note that $\mathrm{Hom}(G,H) \neq \emptyset$ since $\varphi_{\mathrm{triv}} \in \mathrm{Hom}(G,H)$. The set of all automorphisms $\varphi\colon G \to G$ is a subset of $\mathrm{Bi}(G)$, the set of all bijective functions $f\colon G \to G$. We denote the subset of automorphisms of G by $\mathrm{Aut}(G)$. The set $\mathrm{Hom}(G,G)$ is typically denoted by $\mathrm{End}(G)$ and the elements of $\mathrm{End}(G)$ are called **endomorphisms**. The set $\mathrm{End}(G)$ is not a group but instead only a monoid under composition of functions. We will refer to $\mathrm{End}(G)$ as the **endomorphism monoid of G** and $\mathrm{Aut}(G)$ as the **automorphism group of G**.

Definition 26 (Finite/Infinite group)**.** If G is a group and $|G| < \infty$, then we will say that G is a **finite group**. If $|G| = \infty$, then we will say that G is an **infinite group**.

Definition 27 (Exponent of a finite group)**.** If G is a finite group, we define the **exponent of G** to be

$$\mathrm{LCM}\{\mathrm{ord}(g)\ :\ g \in G\}$$

where if $\{\alpha_1,\dots,\alpha_n\} \subset \mathbf{N}$, then $\mathrm{LCM}\{\alpha_1,\dots,\alpha_n\}$ is the least common multiple of $\alpha_1,\dots,\alpha_n$.

3.3 Properties of elements

Definition 28 (Conjugate elements)**.** Given a group G, we say that $g_1,g_2 \in G$ are **conjugate in G** if there exists $h \in G$ such that $g_2 = hg_1h^{-1}$.

Definition 29 (Conjugacy class of an element)**.** Given a group G and an element $g \in G$, the **conjugacy class of g in G** is defined to be the subset

$$[g]_G = \left\{hgh^{-1}\ :\ h \in H\right\}.$$

Definition 30 (Order of an element)**.** Given a group G and $g \in G$, then **order of g** is defined to by $|\langle g \rangle|$ and we denote the order of g by $\mathrm{ord}(g)$. If $\mathrm{ord}(g)$ is finite, we say that g has **finite order**. If $\mathrm{ord}(g)$ is infinite, then we say g has **infinite order**.

A non-trivial finite order element g is sometimes referred to as a **torsion element**. When G has a non-trivial element of finite order, we will say that G has **torsion**. If G has no non-trivial elements of finite order, we will say that G is torsion free.

Definition 31 (Commuting elements)**.** Given a group G, we say $g,h \in G$ **commute** if $gh = hg$.

3.4 Subgroups

Definition 32 (Subgroup)**.** Given a group G and a subset $S \subset G$, we say that S is a **subgroup** if the following conditions hold:

(i) $1 \in S$.

(ii) For each $s_1,s_2 \in S$, we have $s_1s_2 \in S$.

(iii) For each $s \in S$, we have $s^{-1} \in S$.

We S is a subgroup of G, we will write $S \leq G$ in this case.

Remark: A subgroup S of a group is a subset of the group such that the binary operation $\star$ restricts to S to give S a group structure.

Lemma 3.1 (Subgroup Test)**.** *Given a group G, a subset S is a subgroup if and only if $1 \in S$ and for each $s_1, s_2 \in S$, we have $s_1 s_2^{-1} \in S$.*

Definition 33 (Image of a subgroup)**.** Given groups G, H, a homomorphism $\varphi \colon G \to H$, and a subgroup $S \leq G$, we define the **image of S under φ** to be the subgroup of H given by

$$\varphi(S) = \{h \in H \ : \ h = \varphi(s) \text{ for some } s \in S\}.$$

Definition 34 (Pullback of a subgroup)**.** Given groups G, H, a homomorphism $\varphi \colon G \to H$, and a subgroup $T \leq H$, we define the **pullback of T under φ** to be the subgroup of G given by

$$\varphi^{-1}(T) = \{g \in G \ : \ \varphi(g) \in T\}.$$

Lemma 3.2 (Intersection of Subgroups)**.** *Given a group G and a pair of subgroups $S, T \leq G$, the intersection $S \cap T$ is a subgroup of G.*

Definition 35 (Maximal subgroup)**.** Given a group G and a proper subgroup $S \leq G$, we say that S is **maximal in G** if whenever T is a subgroup of G such that

$$S \subset T \subset G,$$

then either $T = S$ or $T = G$.

Definition 36 (Frattini subgroup)**.** Given a group G, let $\mathscr{M}_G$ denote the set of all maximal subgroups of G. We define the **Frattini subgroup** $\mathrm{Frat}(G)$ to be

$$\mathrm{Frat}(G) = \bigcap_{M \in \mathscr{M}_G} M.$$

Definition 37 (Trivial subgroup)**.** Given a group G, the **trivial subgroup of G** is defined to be $\{1\}$.

Definition 38 (Proper subgroup)**.** Given a group and a subgroup $S \leq G$, we say that S is a **proper subgroup** if $S \neq G$.

A non-trivial, proper subgroup of G is a subgroup such that $S \neq G$ and $S \neq \{1\}$.

Definition 39 (Torsion free subgroup)**.** A subgroup $H \leq G$ is said to be **torsion free** if H has no torsion elements.

Definition 40 (Torsion group)**.** A group G is said to be a **torsion group** if every non-trivial element is a torsion element.

Finite groups are torsion groups though there are also infinite torsion groups.

3.5 Centralizers and the center of a group

Definition 41 (Centralizer of an element)**.** Given a group G and $g \in G$, the **centralizer of g in G** is the subgroup

$$C_G(g) = \{h \in G \ : \ gh = hg\}.$$

Definition 42 (Central elements)**.** Given a group G, we say that $g \in G$ is **central** if $C_G(g) = G$.

Definition 43 (Center of a Group)**.** Given a group G, we define the **center of G** to be the subgroup

$$Z(G) = \{g \in G \ : \ g \text{ is central in } G\}.$$

3.6 Conjugation of subgroups

Definition 44 (Conjugate subgroup)**.** Given a group G and a subgroup S of G, the subgroup

$$gSg^{-1} = \left\{gsg^{-1} \ : \ s \in S\right\}$$

is called the g–conjugate of S.

Definition 45. Given a group G and subgroups $S_1, S_2 \leq G$, we say that S_1 is **conjugate in G** to S_2 if there exists $g \in G$ such that

$$S_2 = gS_1g^{-1}.$$

Definition 46 (Normal subgroup)**.** Given a group G and a subgroup $N \leq G$, we say that N is **normal in G** if

$$N = gNg^{-1}$$

for all $g \in G$. When N is a normal subgroup in G, we denote this by $N \lhd G$.

It is a simple matter to see that $\{1\} \lhd G$ and $G \lhd G$ for any group G.

Definition 47 (Non-trivial normal subgroup)**.** Given a group G, we say that a normal subgroup $N \lhd G$ is **non-trivial** if it is proper and non-trivial.

Definition 48 (Simple group)**.** We say that a group G is **simple** if G has no non-trivial normal subgroups.

Definition 49 (Normalizer of a subgroup)**.** Given a group G and a subgroup $H \leq G$, we define the **normalizer of H in G** to be

$$N_G(H) = \left\{g \in G \ : \ gHg^{-1} = H\right\}.$$

Definition 50 (Centralizer of a subgroup)**.** Given a group G and a subgroup $H \leq G$, we define the **centralizer of H in G** to be

$$C_G(H) = \{g \in G \ : \ gh = hg \text{ for all } h \in H\}.$$

Definition 51 (Normal closure)**.** Let G be a group and let $H \leq G$ be a subgroup. The **normal closure** of H in G is defined to be the smallest normal subgroup $N \lhd G$ that contains H. We denote the normal closure of H in G by $\overline{H}$.

Lemma 3.3. *Let G be a group and let $H \leq G$ be a subgroup. For $g \in G$, the following are equivalent:*

(a) $g \in \overline{H}$.

(b) $g \in N$ *for every normal subgroup* $N \lhd G$ *with* $H \leq N$.

Corollary 3.4. *Let G be a group and let $H \leq G$ be a subgroup. Define*

$$\mathrm{Nor}(G,H) = \{N \lhd G \ : \ H \leq G\}.$$

Then

$$\overline{H} = \bigcap_{N \in \mathrm{Nor}(G,H)} N.$$

Definition 52 (Normal core)**.** Let G be a group and let $H \leq G$ be a subgroup. The **normal core** of H in G is defined to be the largest normal subgroup $N \lhd G$ that contains H. We denote the normal closure of H in G by $\mathrm{Core}(H)$.

Lemma 3.5. *Let G be a group and $H \leq G$ a subgroup. For $g \in G$, the following are equivalent:*

(a) $g \in \mathrm{Core}(H)$.

(b) $g \in kHk^{-1}$ *for all* $k \in G$.

Corollary 3.6. *Let* G *be a group and let* $H \leq G$ *be a subgroup. Define Then*

$$\overline{H} = \bigcap_{g \in G} gHg^{-1}.$$

Definition 53 (Characteristic subgroup)**.** Let G be a group and let $H \leq G$ be a subgroup. We say that H is **characteristic** in G if for each automorphism $\psi \in \mathrm{Aut}(G)$, we have $\psi(H) = H$.

Definition 54 (Characteristic core)**.** Let G be a group and let $H \leq G$ be a subgroup. We define the **characteristic core** of H in G to be the largest characteristic subgroup K of G with $K \leq H$. We denote the characteristic core by $\mathrm{Core}_c(H)$

Lemma 3.7. *Let* G *be a group and* $H \leq G$ *a subgroup. For* $g \in G$*, the following are equivalent:*

(a) $g \in \mathrm{Core}_c(H)$.

(b) $g \in \psi(H)$ *for all* $\psi \in \mathrm{Aut}(G)$.

Corollary 3.8. *Let* G *be a group and let* $H \leq G$ *be a subgroup. Then*

$$\mathrm{Core}_c(H) = \bigcap_{\psi \in \mathrm{Aut}(G)} \psi(H).$$

Definition 55 (Fully Invariant)**.** Let G be a group and let $H \leq G$ be a subgroup. We say that H is **fully invariant** in G if for each automorphism $\psi \in \mathrm{End}(G)$, we have $\mathrm{End}(H) \leq H$.

Definition 56 (Malnormal)**.** Let G be a group. We say that a subgroup $H \leq G$ is **malnormal** if $H \cap gHg^{-1} = 1$ for all $g \in G - H$.

3.7 Subgroups generated by subsets and cyclic subgroups

Definition 57 (Subgroup generated by a subset)**.** Given a group G and a subset $X \subset G$, we define the **subgroup generated by X in G** to be the smallest subgroup of G that contains X. We denote this subgroup by $\langle X \rangle$.

Definition 58 (Cyclic subgroup)**.** Given a group G and an element g, the subgroup generated by g is denoted by $\langle g \rangle$ and is called the **cyclic subgroup generated of G by g**.

Lemma 3.9. *If* G *is a group and* $g \in G$*, then* $\langle g \rangle = \{g^n \ : \ n \in \mathbf{Z}\}$.

Lemma 3.10. *Let* G *be a group and let* $S \subset G$*. Then the following are equivalent:*

(a) $g \in \langle S \rangle$.

(b) There exists $s_1, \dots, s_m \in S$ *and* $\varepsilon_1, \dots, \varepsilon_m \in \{\pm 1\}$ *such that*

$$g = s_1^{\varepsilon_1} \dots s_m^{\varepsilon_m}.$$

(c) $g \in H$ *for every subgroup* $H \leq G$ *with* $S \subset H$.

Corollary 3.11. *Let* G *be a group and let* S *be a subset of* G*. Define* $X_S = S \cup S^{-1}$ *where*

$$S^{-1} = \left\{ s^{-1} \ : \ s \in S \right\}.$$

Then $\langle S \rangle$ *is the set of* $g \in G$ *that can be expressed as*

$$g = s_1 \dots s_m$$

for some $s_1, \dots, s_m \in X_S$.

Let $S = \{a, b\}$ where $a, b \in G$ and set $A = a^{-1}$, $B = b^{-1}$. In this case $X_S = \{a, b, A, B\}$. For $g \in G$, we have $g \in \langle S \rangle$ if and only if g can be expressed as some finite product of the elements a, b, A, B. For instance, $aBabAbbbb$ or $abAbAAABa$ are in $\langle S \rangle$. Note that $aA = 1$ and $bB = 1$.

If you want to list off all of the possible elements of $\langle S \rangle$, then you start by selecting one of a, b, A, B. You can stop or pick again. Your next pick can be anything but the inverse of the previous pick. You count to any finite amount, and this will list more and more of $\langle S \rangle$. There is some value to viewing the elements of G that can built with the alphabet X_S (e.g. a, b, A, B when $S = \{a, b\}$) as **words** in this alphabet. The group operation $aA = 1$ provides something like a contraction in the language provided by S (e.g. can not = can't).

Remark: One can view the operation $S \subset G$ get associated to $\langle S \rangle$ as the analog of taking the span of a subset S of some vector space V.

3.8 Subgroup series

This subsection is considered with **subgroup series**.

3.8.1 The commutator subgroup

The **commutator subgroup** is defined by

$$[G, G] = \langle \{[g, h] \ : \ g, h \in G\} \rangle .$$

More generally, given subgroups $H, K \leq G$, we define

$$[H, K] = \langle \{[h, k] \ : \ h \in H, \ k \in K\} \rangle .$$

When G is commutative, the commutator subgroup of G is trivial.

3.8.2 Lower central series

The **lower central series** $\{G_i\}_{i \geq 0}$ of a group G is defined recursively by

$$G_0 = G, \quad G_i = [G, G_i].$$

3.8.3 Derived series

The **derived series** $\left\{G^i\right\}_{i \geq 0}$ of a group G is defined recursively by

$$G^0 = G, \quad G^i = [G^{i-1}, G^{-1}].$$

3.9 Some classes of groups

3.9.1 Nilpotent groups

We say that a group G is nilpotent of step size $j \in \mathbf{N}$ if $G_j = \{1\}$ and $G_i \neq \{1\}$ for all $i \in \mathbf{N}$ with $i < j$.

3.9.2 Solvable groups

We say that a group G is solvable of step size $j \in \mathbf{N}$ if $G^j = \{1\}$ and $G_i \neq \{1\}$ for all $i \in \mathbf{N}$ with $i < j$.

3.9.3 *p*–groups

Given a prime $p \in \mathbf{N}$, we say that G is a ***p*–group** if $|G| = p^n$ for some $n \in \mathbf{N}$.

4 Product sets, quotient groups, and the isomorphism theorems

4.1 Inverse and product sets in groups

4.1.1 Inverse and products sets

Given a group G and subsets $S,T \subset G$, we define the **product set** to be

$$ST = \{st \ : \ s \in S, \ t \in T\}$$

and we define the **inverse set** of S to be

$$S^{-1} = \left\{s^{-1} \ : \ s \in S\right\}.$$

Given three subsets $R,S,T \subset G$, we define

$$RST = \{rst \ : \ r \in R, \ s \in S, \ t \in T\}.$$

We also define the n**th power of** S to be

$$S^n = \{s_1 s_2 \dots s_n \ : \ s_1, s_2, \dots, s_n \in S\}.$$

Note that

$$S^1 = S, \quad S^2 = SS, \quad S^3 = SSS.$$

Below is a list of facts for product and inverse sets.

Lemma 4.1. *Let G be a group and let $R,S,T \subset G$ be subsets of G. Set $\mathscr{I} = \{1\}$*

(a) $R(ST) = (RS)T$.

(b) $(S^{-1})^{-1} = S$ *and* $(S^n)^{-1} = (S^{-1})^n$.

(c) $(ST)^{-1} = T^{-1}S^{-1}$.

(d) $S\mathscr{I} = \mathscr{I}S = S$.

(e) $\emptyset S = S\emptyset = \emptyset$

4.1.2 Monoid structure on the power set $\mathscr{P}$ of G

The product construction above gives a binary operation on the power set $\mathscr{P}(G)$ (i.e. the set of all subsets of G). The elements $\emptyset, \mathscr{I} \in \mathscr{P}(G)$ behavior like 0 and 1 under multiplication. This operation is not a group operation though since inverses rarely exist. In particular, $\mathscr{P}(G)$ is a monoid with this product.

Remark: The reader might not fully understand the above statement that inverses in $\mathscr{P}(G)$ rarely exist. Given S, we do have the inverse set S^{-1}. However, to be an inverse of S under the product operation on $\mathscr{P}(G)$, we need a subset $T \subset G$ such that $ST = \mathscr{I}$. Why does T rarely exist. Take the following general example. Take $S = s_1, s_2$ and take $T = \{t\}$. We see that

$$ST = \{s_1 t, s_2 t\}.$$

If $ST = \mathscr{I}$, then $s_1 t = s_2 t$ since $|ST| = 1$ in this case. However, this implies that $s_1 = s_2$. In particular, if $|S| \geq 2$, then $|ST| \geq 2$ for all $T \neq \emptyset$.

Corollary 4.2. *$S \subset G$ has an inverse in $\mathscr{P}(G)$ if and only if $|S| = 1$.*

The subsets S of G with $|S| = 1$ are singleton sets $\{g\}$. The injective function $G \to \mathscr{P}(G)$ given by $g \mapsto \{g\}$ is a monoid homomorphism and the image of G under this homomorphism is a maximal subgroup. Moreover, the only other subgroup of $\mathscr{P}(G)$ that is not contained in the image of G is the maximal subgroup $\{\emptyset\}$. Here $\{\emptyset\}$ is subset of $\mathscr{P}(G)$ consisting of just the empty set and so $\{\emptyset\} \in \mathscr{P}(\mathscr{P}(G))$.

Given a subset $S \subset G$ and $g \in G$, we define

$$Sg = \{sg \ : \ s \in S\}, \quad gS = \{gs \ : \ s \in S\}.$$

We can combine this construction with the above one with subsets to define subsets of G like $gSg'T$ or $(gS)^{-1}$ or $gSgS^2g^3S^3$ or $(gSg)^3$.

4.1.3 Products with subgroups

When $H \leq G$ is a subgroup, notice that $HH = H$ and $H^{-1} = H$. In fact, we have the following simple observations.

Lemma 4.3. *Let G be a group and let $H \subset G$ be a non-empty subset of G. Then the following are equivalent:*

(a) H is a subgroup of G.

(b) $HH = H$ and $H^{-1} = H$.

Remark: When $H, K \leq G$ are subgroup, the product set HK is not always a subgroup. If either H or K is normal in G, then HK is a subgroup.

4.1.4 Other identities

We see that when $H \leq G$ is a subgroup, then $HH = H$ and $H^{-1} = H$. In order for an element S of $\mathscr{P}(G)$ to be an identity, it is necessary for $SS = S$ and $S^{-1} = S$. Of course, H does not behave like an identity on all of $\mathscr{P}(G)$ when $H \neq \{1\}$ is not the trivial subgroup. We will see that when H is a normal subgroup, the subset of $\mathscr{P}(G)$ of all of the right H–cosets Hg will form a group under this product operation and the trivial coset H will be the identity in this group.

4.2 Cosets of a subgroup

Given a group G, a subgroup $H \leq G$ and $g \in G$, the sets gH and Hg play an important role in understanding the group G. We call these subsets right/left cosets. These sosets gH or Hg of a subgroup H of a group G generalize translates $U + v$ of vector subspaces U of a vector space V. Before discussing cosets in general groups, we will start with a basic but important example. Our group will be $\mathbf{R}^2$ under vector addition and our subgroups will be lines that pass through the origin.

4.2.1 Concrete example: Lines in $\mathbf{R}^2$

In $\mathbf{R}^2$, the vector subspaces of $\mathbf{R}^2$ are the trivial subspace $\{0\}$, $\mathbf{R}^2$ itself, and all lines through the origin. Given a line L through the origin, the points of L can be described by the graph of a function $f(x) = \alpha x$ for some $\alpha \in \mathbf{R}$ (provided L is not the y–axis). The graph of f is the set

$$\text{Graph}(f) = \{(x, f(x)) \ : \ x \in \mathbf{R}\}.$$

In the case $f(x) = \alpha x$ for some fixed real number α, the line L_α is given by the graph of f and is

$$L_\alpha = \{(x, \alpha x) \ : \ x \in \mathbf{R}\}.$$

For $x = 0$, we have $(0,0) \in L_\alpha$. Also, given $x, y \in \mathbf{R}$, we see that

$$(x, \alpha x) - (y, \alpha y) = (x - y, \alpha(x - y)) \in L_\alpha.$$

Hence, L_α is a vector subspace of $\mathbf{R}^2$. Given $v \in \mathbf{R}^2$, we define

$$L_\alpha + v = \{\ell + v \ : \ \ell \in L_\alpha\}.$$

The set $L_\alpha + v$ is a line in $\mathbf{R}^2$ and if $v \neq 0$, it does not pass through $(0,0)$. In particular, $L_\alpha + v$ cannot be a vector subspace of $\mathbf{R}^2$ when $v \neq 0$. Nevertheless, we all appreciate that lines in $\mathbf{R}^2$, whether they pass through the origin or not, have some value.

Remark: We could also take $L_\alpha = \left\{(\alpha^{-1}x, x) \ : \ x \in \mathbf{R}\right\}$ provided $\alpha \neq 0$. Indeed, if $\alpha \neq 0$, then

$$L_\alpha = \{(x, \alpha x) \ : \ x \in \mathbf{R}\} = \left\{(\alpha^{-1}y, y) \ : \ y \in \mathbf{R}\right\}.$$

If $\alpha = 0$, then we have two additional lines:

$$L_x = \{(x, 0) \ : \ x \in \mathbf{R}\}, \quad L_y = \{(0, y) \ : \ y \in \mathbf{R}\}.$$

For simplicity, we will assume $\alpha \neq 0$ in what follows. Everything below also works for the lines L_x, L_y.

An equivalence relationship on $\mathbf{R}^2$ via L_α. We can define a relationship between vectors in $\mathbf{R}^2$ using our line L_α. Specifically, given $v, w \in \mathbf{R}^2$, we write $v \sim_\alpha w$ if $v - w \in L_\alpha$. We will prove that the following holds for all $u, v, w \in \mathbf{R}^2$:

(1) $v \sim_\alpha v$

(2) If $v \sim_\alpha w$, then $w \sim_\alpha v$.

(3) If $v \sim_\alpha w$ and $w \sim_\alpha u$, then $v \sim_\alpha u$.

For (1), $v \sim_\alpha v$ if $v - v \in L_\alpha$. Well, $v - v = (0,0) \in L_\alpha$ for $x = 0$.

For (2), if $v \sim_\alpha w$, then $v - w \in L_\alpha$. In particular, $v - w = (x, \alpha x)$ for some $x \in \mathbf{R}$. Of course,

$$w - v = -(v - w) = -(x, \alpha x) = (-x, \alpha(-x)) \in L_\alpha.$$

Hence $w \sim_\alpha v$.

For (3), if $v \sim_\alpha w$ and $w \sim_\alpha v$, then

$$v - w = (x, \alpha x), \quad w - u = (y, \alpha y).$$

Now,

$$\begin{aligned} v - u &= v + (-w + w) - u \\ &= v - w + w - u \\ &= (v - w) + (w - u) = (x, \alpha x) + (y, \alpha y) \\ &= (x + y, \alpha(x + y)) \in L_\alpha. \end{aligned}$$

Hence, $v \sim_\alpha u$.

The equivalence class of a vector. Given $v \in \mathbf{R}^2$, we define

$$[v]_{L_\alpha} = \left\{ w \in \mathbf{R}^2 \ : \ v \sim_\alpha w \right\}.$$

We call this the **equivalence of** v **under** $\sim_\alpha$. We know want to determine what exactly $[v]_{L_\alpha}$ is (e.g. what types of vectors w are equivalent to v under $\sim_\alpha$). We will prove that

$$[v]_{L_\alpha} = L_\alpha + v.$$

We begin by first studying when $v \sim_\alpha w$. We know that $v = (v_1, v_2)$ for some $v_1, v_2 \in \mathbf{R}$. Now, assume that $w = (w_1, w_2)$ and $v \sim_\alpha w$. Then $v - w \in L_\alpha$. In particular,

$$v - w = (x, \alpha x)$$

for some $x \in \mathbf{R}$. On the other hand,

$$v - w = (v_1, v_2) - (w_1, w_2) = (v_1 - w_1, v_2 - w_2).$$

Hence

$$(v_1 - w_1, v_2 - w_2) = (x, \alpha x).$$

Therefore,

$$v_1 - w_1 = x, \quad v_2 - w_2 = \alpha x.$$

Solving for w_1, w_2, we obtain

$$w_1 = v_1 - x, \quad w_2 = v_2 - \alpha x,$$

and so

$$w = (v_1 - x, v_2 - \alpha x).$$

Now, if $w = (v_1 - x, v_2 - \alpha x)$ for some $x \in \mathbf{R}$, we will check that $v \sim_\alpha w$. To that end, we have

$$v - w = (v_1, v_2) - (v_1 + x, v_2 + \alpha x) = (v_1 - v_1 + x, v_2 - v_2 + \alpha x) = (x, \alpha x) \in L_\alpha.$$

Remark: Above, we showed that $v \sim_\alpha w$ if and only if $w = (v_1 - x, v_2 - \alpha x)$ for some $x \in \mathbf{R}$. This condition on w is also equivalent to

$$w = (v_1 + x, v_2 + \alpha x)$$

for some $x \in \mathbf{R}$. To see this, simply note that

$$(v_1 - x, v_2 - \alpha x) = (v_1 + (-x), v_2 + \alpha(-x)).$$

In particular, if $x \in \mathbf{R}$ satisfies

$$w = (v_1 - x, v_2, -\alpha x),$$

then $-x = y \in \mathbf{R}^2$ satisfies

$$w = (v_1 + y, v_2 + \alpha y).$$

Summary: If $v = (v_1, v_2)$, $w = (w_1, w_2)$, then the following are equivalent:

(2) $w = (v_1 - x, v_2 - \alpha x)$ for some $x \in \mathbf{R}$.

(3) $w = (v_1 + x, v_2 + \alpha x)$ for some $x \in \mathbf{R}$.

(4) $v \sim_\alpha w$.

The set $L_\alpha + v$. Next, we consider when $w \in L_\alpha + v$. As before, set $v = (v_1, v_2)$ and $w = (w_1, w_2)$. If $w \in L_\alpha + v$, then

$$w = u + v$$

for some $u \in L$. In coordinates, this becomes

$$(w_1, w_2) = (x, \alpha x) + (v_1, v_2).$$

Hence

$$w_1 = x + v_1, \quad w_2 = \alpha x + v_2.$$

Conversely, if

$$w = (x + v_1, \alpha x + v_2)$$

for some $x \in R$, then

$$w = (x + v_1, \alpha x + v_2) = (x, \alpha x) + (v_1, v_2) = u + v$$

for some $u \in L_\alpha$.

Remark: As before, if $w = (x + v_1, \alpha x + v_2)$ for some $x \in \mathbf{R}$, then $w = (v_1 - y, v_2 - \alpha y)$ for some $y \in \mathbf{R}$

Summary: If $v = (v_1, v_2)$, $w = (w_1, w_2)$ and $w \in L_\alpha + v$, then the following are equivalent:

(1) $w \in L_\alpha + v$.

(2) $w = (v_1 - x, v_2 - \alpha x)$ for some $x \in \mathbf{R}$.

(3) $w = (v_1 + x, v_2 + \alpha x)$ for some $x \in \mathbf{R}$.

(4) $v \sim_\alpha w$.

Corollary 4.4. *$L_\alpha + v = [v]_{L_\alpha}$.*

The set of equivalence classes. Let $V_{L_\alpha} \subset \mathscr{P}(\mathbf{R}^2)$ be the set of subsets of $\mathbf{R}^2$ given by

$$[v]_{L_\alpha} = \left\{ w \in \mathbf{R}^2 \ : \ v \sim_\alpha w \right\}.$$

We note that given $v, v' \in \mathbf{R}^2$, either $[v]_{L_\alpha} = [v']_{L_\alpha}$ of $[v]_{L_\alpha} \cap [v']_{L_\alpha} = \emptyset$. In particular, every vector v is contained in a unique element of V_{L_α}, namely $[v]_{L_\alpha} \in VL_\alpha$.

Remark: The set V_{L_α} is the set of all distinct subsets $[v]_{L_\alpha}$ of $\mathbf{R}^2$ as we vary over all $v \in \mathbf{R}^2$. If $v \sim_\alpha w$, then $[v]_{L_\alpha} = [w]_{L_\alpha}$ represent that same point (we think of the sets of V_{L_α} as points) in V_{L_α}. This can be a little difficult to think about at first. However, do remember that the sets $[v]_{L_\alpha}$ are just translates of the line L_α by v in $\mathbf{R}^2$. In particular, V_{L_α} is the set of all lines that are parallel to L_α including L_α itself.

There is a natural function $q_{L_\alpha} \colon \mathbf{R}^2 \to V_{L_\alpha}$ defined by $q_{L_\alpha}(v) = [v]_{L_\alpha}$. The function q_{L_α} is surjective but not injective. Viewing V_{L_α} as the set of all lines that are parallel to L_α, the function q_{L_α} sends the vector v to the unique line in V_{L_α} that contains v. We also see that

$$q_{L_\alpha}^{-1}([v]_{L_\alpha}) = L_\alpha + v \subset \mathbf{R}^2.$$

Finally, notice that $[0]_{L_\alpha} = L_\alpha$

A natural vector space structure on $V_{L\alpha}$. We will now endow $V_{L\alpha}$ with a real vector space structure such that $q_{L\alpha}\colon \mathbf{R}^2 \to V_{L\alpha}$ is a linear function. Given $[v]_{L\alpha}, [w]_{L\alpha} \in V_{L\alpha}$ and $\lambda \in \mathbf{R}$, we define

$$[v]_{L\alpha} + [w]_{L\alpha} = [v+w]_{L\alpha}, \quad \lambda[v]_{L\alpha} = [\lambda v]_{L\alpha}.$$

These two operations will be our vector addition and scalar multiplication on $V_{L\alpha}$. The zero vector on $V_{L\alpha}$ with be $[0]_{L\alpha}$.

Now, the definitions for addition and scalar multiplication might not be well defined. If $[v]_{L\alpha} = [v']_{L\alpha}$ and $[w]_{L\alpha} = [w']_{L\alpha}$, we need to check that

$$[v+w]_{L\alpha} = [v'+w']_{L\alpha}.$$

We also must check that

$$[\lambda v]_L L_\alpha = [\lambda v']_{L\alpha}.$$

For the first verification, since $v \sim_\alpha v'$ and $w \sim_\alpha w'$, there are $s,t \in \mathbf{R}$ such that

$$v' = (v_1+s, v_2+\alpha s), \quad w' = (w_1+t, w_2+\alpha t)$$

where $v = (v_1,v_2)$ and $w = (w_1,w_2)$. From this, we have

$$\begin{aligned} v'+w' &= (v_1+s, v_2+\alpha s) + (w_1+t, w_2+\alpha t) \\ &= (v_1+w_1, v_2+w_2) + (s+t, \alpha(s+t)) \\ &= v+w+u \end{aligned}$$

where $u \in L$. Hence $(v+w) \sim_\alpha (v'+w')$ and so

$$[v+w]_{L\alpha} = [v'+w']_{L\alpha}.$$

For the second verification, we have

$$\begin{aligned} \lambda v' &= \lambda(v_1+s, v_2+\alpha s) \\ &= (\lambda v_1 + \lambda s, \lambda v_2 + \alpha\lambda s) \\ &= (\lambda v_1, \lambda v_2) + (\lambda s, \alpha\lambda s) \\ &= \lambda v + u \end{aligned}$$

where $u \in L$. Hence $\lambda v \sim_\alpha \lambda v'$ and so

$$[\lambda v]_{L\alpha} = [\lambda v']_{L\alpha}.$$

We need to check the following:

- $[v]_{L\alpha} + [w]_{L\alpha} = [w]_{L\alpha} + [v]_{L\alpha}$.
- $[v]_{L\alpha} + ([w]_{L\alpha} + [u]_{L\alpha}) = ([v]_{L\alpha} + [w]_{L\alpha}) + [u]_{L\alpha}$.
- $[v]_{L\alpha} + [0]_{L\alpha} = [v]_{L\alpha}$.
- $[v]_{L\alpha} + [-v]_{L\alpha} = [0]_{L\alpha}$.
- $\lambda([v]_{L\alpha} + [w]_{L\alpha}) = \lambda[v]_{L\alpha} + \lambda[w]_{L\alpha}$.
- $(\lambda+\lambda')[v]_{L\alpha} = \lambda[v]_{L\alpha} + \lambda'[v]_{L\alpha}$.
- $(\lambda\lambda')[v]_{L\alpha} = \lambda(\lambda'[v]_{L\alpha})$.
- $1[v]_{L\alpha} = [v]_{L\alpha}$.

We leave the reader to verify that the above equalities hold.

The quotient map. We will now show that $q_{L_\alpha} \colon \mathbf{R}^2 \to V_{L_\alpha}$ is a linear function. Recall that $q_{L_\alpha}(v) = [v]_{L_\alpha}$. Given $v, w \in \mathbf{R}$ and $\lambda, \lambda' \in \mathbf{R}$, we have

$$\begin{aligned} q_{L_\alpha}(\lambda v + \lambda' w) &= [\lambda v + \lambda' w]_{L_\alpha} \\ &= [\lambda v]_{L_\alpha} + [\lambda' w]_{L_\alpha} \\ &= \lambda [v]_{L_\alpha} + \lambda' [w]_{L_\alpha} \\ &= \lambda q_{L_\alpha}(v) + \lambda' q_{L_\alpha}(w). \end{aligned}$$

Hence, q_{L_α} is a linear function.

Corollary 4.5. *Let $L_\alpha = \{(x, \alpha x) \ : \ x \in \mathbf{R}\}$ and V_{L_α} be the set of equivalence classes $\{[v]_{L_\alpha}\}$ with the vector space structure defined above, and $q_{L_\alpha} \colon \mathbf{R}^2 \to V_{L_\alpha}$ be given by $q_{L_\alpha}(v) = [v]_{L_\alpha}$.*

(a) *q_{L_α} is a surjective linear function.*

(b) $\ker(q_{L_\alpha}) = L_\alpha$.

The vector space V_{L_α} is called the **quotient space of $\mathbf{R}^2$ by L_α**. The function q_{L_α} is called the **canonical projection** or the **quotient map**.

An explicit example of this construction. As way of example and also for completeness, we will consider the above when the line is either L_x or L_y.

First, we consider the construction for the line L_x. For $v, w \in \mathbf{R}^2$ with $v = (v_1, v_2)$ and $w = (w_1, w_2)$, we see that $v \sim_x w$ if and only if

$$(v_1 - w_1, v_2 - w_2) \in L_x = \{(x, 0) \ : \ x \in \mathbf{R}\}.$$

In particular, $v \sim_x w$ if and only if $v_2 = w_2$. In particular, every $v \in \mathbf{R}^2$ is equivalent to a vector of the form $(0, y)$ where $y = v_2$. Setting V_x to be the set of equivalence classes $[v]_x$, we can view V_x as

$$V_x = \{[(0, y)]_x \ : \ y \in \mathbf{R}\}.$$

The vector space operations are given by

$$[(0, y_1)]_x + [(0, y_2)]_x = [(0, y_1 + y_2)]_x, \quad \lambda [(0, y)]_x = [(0, \lambda y)]_x.$$

Finally, the quotient map $q_x \colon \mathbf{R}^2 \to V_x$ is given by $q_x((x, y)) = [(0, y)]_x$. It is not hard to see that if we remove some of the overlying notation, the function $q_x((x, y)) = y$. In particular, q_x is the projection map onto the second coordinate. In fact, it is better in this case to think of $q_x \colon \mathbf{R}^2 \to L_y$ since

$$L_y = \{(0, y) \ : \ y \in \mathbf{R}\}.$$

The reader can check that though q_x is not injective on $\mathbf{R}^2$, the restriction of this function to L_y is bijective. Of course, $\ker(q_x) = L_x$. Setting $\iota_x \colon L_x \to \mathbf{R}^2$ to be $\iota((x, 0)) = (x, 0)$, we have

$$\begin{tikzcd} L_x \arrow[r, "\iota_x"] & \mathbf{R}^2 \arrow[r, "q_x"] & L_y \end{tikzcd}$$

where $\iota_x(L_x) = \ker(q_x)$. This is an example of a short exact sequence of vector spaces. Taking $\iota_\alpha \colon L_\alpha \to \mathbf{R}^2$ to be $\iota_\alpha(v) = v$, we have

$$\begin{tikzcd} L_\alpha \arrow[r, "\iota_\alpha"] & \mathbf{R}^2 \arrow[r, "q_{L_\alpha}"] & V_{L_\alpha} \end{tikzcd}$$

where $\iota_\alpha(L_\alpha) = \ker(q_{L_\alpha})$. When the line is L_y, we have $q_y((x, y)) = x$ and

$$\begin{tikzcd} L_y \arrow[r, "\iota_y"] & \mathbf{R}^2 \arrow[r, "q_y"] & L_x \end{tikzcd}$$

where $\iota_y(L_y) = \ker(q_y)$.

4.2.2 Cosets

Given a group G, a subgroup $H \leq G$, and $g \in G$, the **left/right H–cosets associated to g** are

$$gH = \{gh \ : \ h \in H\}, \quad Hg = \{hg \ : \ h \in H\}.$$

If $g \notin H$, neither gH or Hg are subgroups of G since $1 \notin gH, Hg$. The sets gH and Hg are analogous to the lines in $\mathbf{R}^2$ that do not pass through the origin.

Lemma 4.6. *Let G be a group, $H \leq G$ a subgroup, and $g \in G$. Then $gH = (gHg^{-1})g$. In particular, every left H–coset gH is equal to a right coset of a conjugate of H, namely $K = gHg^{-1}$.*

Proof. For this, we have $gH = (gH)(g^{-1}g) = (gHg^{-1})g$. □

4.2.3 Normality of subgroups via coset conditions

The following gives an equivalent condition for a subgroup $N \leq G$ to be normal in G in terms of left and right cosets.

Lemma 4.7. *Let G be a group and $N \leq G$ a subgroup. The following are equivalent:*

(a) $N \lhd G$.

(b) $gN = Ng$ for all $g \in G$.

Proof. For (a) implies (b), since $N \lhd G$, we know that $gNg^{-1} = N$ for all $g \in G$. In particular

$$gN = gNg^{-1}g = Ng$$

for all $g \in G$. Hence N satisfies (b). For (b) implies (a), we know that $gN = Ng$ for all $g \in G$. In particular

$$gNg^{-1} = Ngg^{-1} = N,$$

and so N satisfies (a). □

4.3 Coset and quotient spaces

We will now undertake the construction of a quotient space for each subgroup of a group. This will be done via an equivalence relation. The construction is analogous to what we did above for lines in $\mathbf{R}^2$.

Warning: Though our construction here is identical to that done for lines in $\mathbf{R}^2$, the set of equivalence classes will not always have a natural group structure (i.e. a group structure such that the quotient map is a homomorphism). The subgroups with quotient sets that have natural group structures are the normal subgroups. The normality is needed in proving that the group operation is well defined.

4.3.1 Equivalence relations via subgroups

Given a group G and a subgroup $H \leq G$, we define an equivalence relation $\sim_H$ on G. Specifically, for $g, k \in G$, we define $g \sim_H k$ when $gk^{-1} \in H$. We must verify that the following properties hold for all $g, k, \ell \in G$:

(1) $g \sim_H g$.

(2) If $g \sim_H k$, then $k \sim_H g$.

(3) If $g \sim_H k$ and $k \sim_H \ell$, then $g \sim_H \ell$.

For (1), simply note that $gg^{-1} = 1 \in H$ since H is a subgroup. For (2), if $gk^{-1} \in H$, since H is a subgroup, we have $(gk^{-1})^{-1} \in H$. However,

$$(gk^{-1})^{-1} = (k^{-1})^{-1}g^{-1} = kg^{-1} \in H.$$

Hence, $k \sim_H g$. For (3), if $g \sim_H k$ and $k \sim_H \ell$, then $gk^{-1} \in H$ and $k\ell^{-1} \in H$. Since H is a subgroup, we have $(gk^{-1})(k\ell^{-1}) \in H$. However,

$$(gk^{-1})(k\ell^{-1}) = g(k^{-1}k)\ell^{-1} = g\ell^{-1} \in H.$$

Hence $g \sim_H \ell$.

Remark: In verifying that $\sim_H$ satisfies (1), (2), and (3), we had to use:

(1') $1 \in H$.

(2') If $h \in H$, then $h^{-1} \in H$.

(3') If $h, h' \in H$, then $hh' \in H$.

(1'), (2'), and (3') are equivalent to H being a subgroup.

We can define another equivalence relation $\approx_H$ on G from H. Given $g, k \in G$, we define $g \approx_H k$ if $g^{-1}k \in H$. We must verify that the following properties hold for all $g, k, \ell \in G$:

(1) $g \approx_H g$.

(2) If $g \approx_H k$, then $k \approx_H g$.

(3) If $g \approx_H k$ and $k \approx_H \ell$, then $g \approx_H \ell$.

For (1), simply note that $g^{-1}g = 1 \in H$ since H is a subgroup. For (2), if $g \approx_H k$, then $g^{-1}k \in H$. Since H is a subgroup, we have $(g^{-1}k)^{-1} = k^{-1}g \in H$ and hence $k \approx_H g$. For (3), if $g \approx_H k$ and $k \approx_H \ell$, then $g^{-1}k, k^{-1}\ell \in H$. Since H is a subgroup, we have $g^{-1}kk^{-1}\ell = g^{-1}\ell \in H$.

4.3.2 Equivalence classes associated subgroups

Given $g \in G$, we define the **right H–equivalence class of** g to be

$$[g]_H = \{k \in G \;:\; g \sim_H k\}.$$

Given $k \in [g]_H$, we know that $gk^{-1} \in H$ and so $gk^{-1} = h$ for some $h \in H$. Solving for k, we have $k = h^{-1}g \in Hg$. Conversely, if $k \in Hg$, then $k = hg$ for some $h \in H$. In particular, $h^{-1} = gk^{-1} \in H$ and so $g \sim_H k$.

Lemma 4.8. *Let G be a group, $H \leq G$ a subgroup, and $g \in G$. Then $[g]_H = Hg$. That is, the right H–equivalence class $[g]_H$ of g is equal to the right coset Hg. Also, for each $k \in G$, we have either $Hg = Hk$ or $Hg \cap Hk = \emptyset$.*

Given $g \in G$, we define the **left H–equivalence class of** g to be

$$_H[g] = \{k \in G \;:\; g \approx_H k\}.$$

Given $k \in {}_H[g]$, we know that $g^{-1}k \in H$. In particular, $g^{-1}k = h$ for some $h \in H$. Solving for k, we find $k = gh$ for some $h \in H$. Converse, if $k = gh$ for some $h \in H$, solving for h, we find $g^{-1}k = h \in H$.

Lemma 4.9. *Let G be a group, $H \leq G$ a subgroup, and $g \in G$. Then ${}_H[g] = gH$. That is, the left H–equivalence class ${}_H[g]$ of g is equal to the left coset gH. Also, for each $k \in G$, we have either $gH = kH$ or $gH \cap kH = \emptyset$.*

4.3.3 Quotient sets and quotient maps

We define Q_H to be the set of distinct equivalence classes $[g]_H$ = Hg. We have a surjective function

$$\varphi_H \colon G \longrightarrow Q_H$$

defined by

$$\varphi_H(g) = [g]_H = Hg.$$

We call Q_H the **right quotient space of** G **by** H and we call φ_H the **quotient map associated to** H. We will denote Q_H usually by $H\backslash G$. The set $H\backslash G$ is also called the **right** H**–coset space of** G.

We define ${}_HQ$ to be the set of distinct equivalence classes ${}_H[g] = gH$. We have a surjective function

$${}_H\varphi \colon G \to {}_HQ$$

defined by

$${}_H\varphi(g) = {}_H[g] = gH.$$

We call H_Q the **left quotient space of** G **by** H and call ${}_H\varphi$ the **quotient map associated to** H. We will denote ${}_HQ$ usually by G/H. The set G/H is also called the **left** H**–coset space of** G. In this notation, we have

$$\varphi_H \colon G \to G/H, \quad \varphi_H(g) = Hg$$

and

$${}_H\varphi_H \colon G \to H\backslash G, \quad \varphi_H(g) = gH.$$

4.3.4 Binary operations on cosets spaces

We will now try to define a group structure of $Q_H = G/H$. The elements of G/H are the distinct right H–cosets Hg. These sets partition G into a disjoint family of subsets, namely the right cosets.

Remark: Returning to our example when $G = \mathbf{R}^2$ and L is a line through the origin, we see that L is a subgroup of $\mathbf{R}^2$ under vector addition. The right cosets of L in $\mathbf{R}^2$ are $L+x$ and are lines that that are parallel to L. In the case $L = L_x$ is the x–axis, the cosets are the horizontal lines $L_{x,y_0} = \{(t,y_0) \ : \ t \in \mathbf{R}\}$ where $y_0 \in \mathbf{R}$. Alternatively, this is the horizontal line that intersects the y axis at $(0,y_0)$. These cosets partition $\mathbf{R}^2$ into a family of lines. Each point $v = (x_0,y_0) \in \mathbf{R}^2$ is contained in exactly one of these lines, namely the family of horizontal line. In the coset space, each of these lines horizontal lines in $\mathbf{R}^2$ is mapped to a single point in $\mathbf{R}^2/L$. Indeed, we mapped the horizontal line that meets the y–axis at $(0,y_0)$ to (essentially) y_0. Technically, we map this line to the coset $L_x + (0,y_0)$. This is just projection onto the y–axis. We send each vector to its y–coordinate value. When thinking about G/H as a set, think of G as $\mathbf{R}^2$ and G/H as the y–axis. This concrete model for the abstract concept G/H can be helpful for intuition.

Now, G/H is the set of distinct right H–cosets of G. We want to define a group structure on G/H such that φ_H is a group homomorphism. In particular, if $\cdot \colon G/H \times G/H \to G/H$ is the group operation, then

$$Hg \cdot Hg' = \varphi_H(g) \cdot \varphi_H(g) = \varphi_H(gg') = H(gg').$$

We will now consider what precisely this means. Before doing so, we will return again to the example of a line in $\mathbf{R}^2$. We defined the group operation of $\mathbf{R}^2/L$ to be $(L+v)\cdot(L+v') = L+(v+v')$. Of course, we think of $\cdot$ here as an additive operation in this case. Nevertheless, for this group operation to be well defined, we need to know that if we take vectors $w \in L+v$ and $w' \in L+v'$, that $L+(w+w') = L+(v+v')$. Put another way, this operation is well defined when

$$(L+v)+(L+v') = L+(v+v')$$

where

$$\begin{aligned} L+v &= \left\{u \in \mathbf{R}^2 \ : \ u = \ell + v \text{ for some } \ell \in L\right\} \\ L+v' &= \left\{u \in \mathbf{R}^2 \ : \ u = \ell + v' \text{ for some } \ell \in L\right\} \\ L+(v+v') &= \left\{u \in \mathbf{R}^2 \ : \ u = \ell + v + v' \text{ for some } \ell \in L\right\} \\ (L+v)+(L+v') &= \left\{u \in \mathbf{R}^2 \ : \ u = \ell + v + \ell' + v' \text{ for some } \ell, \ell' \in L\right\}. \end{aligned}$$

Here the vectors v, v' are fixed but could be any pair of vectors. That

$$L+(v+v') = (L+v)+(L+v')$$

follows from commutativity of vector addition and that L is a vector subspace. For instance, since L is a vector subspace, then for any $\ell, \ell' \in L$, we know that $\ell + \ell' = \ell''$ for some $\ell'' \in L$. Hence,

$$\begin{aligned} (L+v)+(L+v') &= \left\{u \in \mathbf{R}^2 \ : \ u = \ell + v + \ell' + v' \text{ for some } \ell, \ell' \in L\right\} \\ &= \left\{u \in \mathbf{R}^2 \ : \ u = \ell + \ell' + v + v' \text{ for some } \ell, \ell' \in L\right\} \\ &= \left\{u \in \mathbf{R}^2 \ : \ u = \ell'' + v + v' \text{ for some } \ell'' \in L\right\} \\ &= L+(v+v'). \end{aligned}$$

Returning to our attempt to give G/H a natural group structure (e.g. so that φ_H is a group homomorphism), we require

$$HgHg' = Hgg'$$

where

$$\begin{aligned} Hg &= \{hg \ : \ h \in H\} \\ Hg' &= \left\{hg' \ : \ h \in H\right\} \\ Hgg' &= \left\{hgg' \ : \ h \in H\right\} \\ HgHg' &= \left\{hgh'g' \ : \ h, h' \in H\right\}. \end{aligned}$$

In particular, we require

$$HgHg' = \left\{hgh'g' \ : \ h, h' \in H\right\} = \left\{hgg' \ : \ h \in H\right\} = Hgg'.$$

To begin, *this is not always true*.

Lemma 4.10. *If G is a commutative group and $H \leq G$ is a subgroup, then $HgHg' = Hgg'$ for all $g, g' \in G$.*

Proof. Using the commutativity of G and that H is a subgroup, we have

$$\begin{aligned} HgHg' &= \left\{hgh'g' \ : \ h, h' \in H\right\} \\ &= \left\{hh'gg' \ : \ h, h' \in H\right\} \\ &= \left\{h''gg' \ : \ h'' \in H\right\} = Hgg'. \end{aligned}$$

□

Lemma 4.11. *Let G be a group and $H \leq G$ a subgroup. If $gH = Hg$ for all $g \in G$, then $HgHg' = Hgg'$ for all $g, g' \in G$.*

Proof. We must prove that $HgHg' = Hgg'$ for all $g, g' \in G$. First, we will give a proof using the set condition that $gH = Hg$. We will prove that $HgHg' = Hgg'$ using product construction and some of the observations on that topic (e.g. $HH = H$ when H is a subgroup). To this end, we have

$$HgHg' = H(gH)g' = H(Hg)g' = HHgg' = (HH)gg' = Hgg'.$$

Second, we will give a more concrete argument that works on the level of points. We will prove the sets are equal using the definition of set equality. We will prove that $HgHg' \subset Hgg'$ and $Hgg' \subset HgHg'$ for all $g, g' \in G$. To prove the first inclusion, we must prove for each $x \in HgHg'$, that $x \in Hgg'$. Since $x \in HgHg'$, we know

$$x = hgh'g' \tag{1}$$

for some $h, h' \in H$. Since $gh' \in gH$ and $gH = Hg$, we know that there exists $h''' \in H$ such that

$$gh' = h'''g. \tag{2}$$

Using (1) and (2), we obtain

$$x = hh'''gg'. \tag{3}$$

Since H is a subgroup, we know $hh''' = h'''' \in H$ and so using (3), we see that $x = h''''gg' \in Hgg'$. This completes the proof of the first inclusion.

To prove the second inclusion, we must prove that for each $x \in Hgg'$, that $x \in HgHg'$. Since $x \in Hgg'$, we know

$$x = hgg' \tag{4}$$

for some $h \in H$. Since $gH = Hg$, we know that

$$hg = gh'' \tag{5}$$

for some $h'' \in H$. Using (4) and (5), we obtain

$$x = gh''g' \tag{6}$$

for some $h'' \in H$. Now,

$$x = (h_0 h_0^{-1})gh''g' \tag{7}$$

for any $h_0 \in H$. Using (6), (7), and that $1 \in H$, we see $x \in HgHg'$. □

4.3.5 Normal subgroups and quotient groups

Let G be a group and let $H \lhd G$ be a normal subgroup. As defined above, we have the space G/H of right H–cosets. Given $Hg_1, Hg_2 \in G/H$, we define the group operation $\cdot$ on G/H by

$$Hg_1 \cdot Hg_2 = Hg_1g_2. \tag{8}$$

Since H is normal, we know that $gH = Hg$ for all $g \in G$ and so from our work above, we know that

$$(Hg_1)(Hg_2) = Hg_1g_2.$$

Also, as H is normal, we know that $gH = Hg$ for all $g \in G$. Hence,

$$(g_1H)(g_2H) = (Hg_1)(Hg_2) = Hg_1g_2 = g_1g_2H.$$

Consequently, we define the group operation on $H \backslash G$ to be

$$g_1H \cdot g_2H = g_1g_2H. \tag{9}$$

We will now verify that G/H is a group with this operation. The identity is the trivial coset H and the inverse of Hg is Hg^{-1}. That H is the identity can be seen via:

$$HHg = Hg, \quad HgH = HHg = Hg.$$

For inverses, we see that

$$HgHg^{-1} = Hgg^{-1} = H, \quad Hg^{-1}Hg = Hg^{-1}g = H.$$

That the operation is associative follows from the associativity of multiplication on G.

Since $gH = Hg$ for all $g \in G$, we see that $H\backslash G$ and G/H are isomorphic groups. We will denote this group typically by G/H and will work with right H–cosets. The group G/H is called the **quotient group of G by H**.

We next verify that $\varphi_H \colon G \to G/H$ is a homomorphism. For that, simply note that $\varphi_H(g) = Hg$, and

$$\varphi_H(gg') = Hgg' = HgHg' = \varphi_H(g)\varphi_H(g').$$

We also see that

$$\ker(\varphi_H) = H.$$

4.3.6 Summary of quotients

Given a group G and a subgroup H, we have the following:

- A set G/H which is the set of all distinct right cosets Hg. This is called the **quotient space of G by H**.
- A surjective function $\varphi_H \colon G \to G/H$ defined by $\varphi_H(g) = Hg$. This is called the **quotient map**.
- When H is a normal subgroup, G/H is a group under the multiplication $Hg \cdot Hg' = H(gg')$. We denote this multiplication operation simply by $HgHg'$ and note that $HgHg' = Hgg'$ as sets.
- When H is a normal subgroup, the identity in G/H with this multiplication is the **trivial coset** $H \in G/H$. The subgroup H is equivalence class of all elements of G that are equivalent to some element of H. That is, $H = Hh$ for all $h \in H$ and $H \neq Hg$ for all $g \notin H$.
- When H is a normal subgroup, φ_H is a surjective group homomorphism. It is also sometimes referred to in this case as the **canonical homomorphism** or **canonical epimorphism**.

4.3.7 Index of a subgroup

Definition 59 (Index of a subgroup)**.** Given a group G with subgroup $H \leq G$, we define the **index of H in G** to be $[G:H] = |G/H|$.

Since $gH = Hg^{-1}$ for all $g \in G$ and subgroups $H \leq G$, it follows that $|H\backslash G| = |G/H|$. Additionally, some authors use $|G:H|$ to denote the index of H.

Definition 60 (Finite index subgroup)**.** Given a group G and a subgroup $H \leq G$, we say H **has finite in** G if $[G:H] < \infty$.

When $H \lhd G$ is a finite index normal subgroup, G/H is a finite group.

Lemma 4.12. *If G is a finite group and $H \leq G$, then*

$$|G| = |H|\,[G:H].$$

This lemma follows easily from the following basic fact: If $f\colon X \to Y$ is a surjective function and X,Y are finite sets, then

$$|X| = \sum_{y \in Y} \left| f^{-1}(y) \right|.$$

For the lemma, we take f to be the quotient map $\varphi_H\colon G \to G/H$.

4.4 The Isomorphism Theorems

The isomorphism theorem for groups are a collection of results that hold for all groups and relate group homomorphisms, subgroups, and quotient groups.

4.4.1 The First Isomorphism Theorem

Let G,H be groups and let $\varphi\colon G \to H$ be a homomorphism with $K = \ker(\varphi)$ and $L = \varphi(G)$. Since K is normal, we have a quotient group G/K and surjective homomorphism $\varphi_K\colon G \to G/K$. In total, we have

$$\begin{tikzcd}
G \arrow[rr, "\varphi"] \arrow[dr, "\varphi_K"'] & & L = \varphi(H) \leq H \\
& G/K &
\end{tikzcd}$$

We define a function $\psi\colon G/K \to L \leq H$ defined by $\psi(Kg) = \varphi(g)$. By definition of ψ, we have

$$\begin{tikzcd}
G \arrow[rr, "\varphi"] \arrow[dr, "\varphi_K"'] & & L = \varphi(H) \leq H \\
& G/K \arrow[ur, "\psi"'] &
\end{tikzcd}$$

with $\psi \circ \varphi_K = \varphi$. Since φ_K and φ are homomorphisms, we have

$$\varphi_K(gg') = \varphi_K(g)\varphi_K(g'), \quad \varphi(gg') = \varphi(g)\varphi(g')$$

for all $g,g' \in G$. Hence,

$$\psi(\varphi_K(gg')) = \psi(\varphi_K(g)\varphi_K(g')) = \varphi(gg') = \varphi(g)\varphi(g')$$

for all $g,g' \in G$. In particular, for each $Kg, Kg' \in G/K$, we have

$$\psi(KgKg') = \psi(Kgg') = \psi(\varphi_K(gg')) = \varphi(g)\varphi(g') = \psi(Kg)\psi(Kg').$$

Thus, ψ is a homomorphism.

Remark: If we have three groups G,H,K and homomorphisms $\psi\colon G \to H$ and $\varphi\colon G \to K$, then if $\phi\colon K \to H$ is a function such that $\psi = \phi \circ \varphi$, then ϕ is also a homomorphism. That is, we have the diagram

$$\begin{tikzcd}
G \arrow[rr, "\varphi"] \arrow[dr, "\psi"'] & & H \\
& K \arrow[ur, "\phi"'] &
\end{tikzcd}$$

with ψ,φ homomorphisms and $\psi = \phi \circ \varphi$, then ϕ is a homomorphism. We leave this for the reader to verify.

Returning to our main setting, we will now prove that $\psi\colon G/K \to L$ is an isomorphism. We have already verified that ψ is a homomorphism and so it remains to check that ψ is a bijection.

First, that ψ is onto follows from the fact that $\varphi\colon G \to L \leq H$ is onto and $\varphi = \psi \circ \varphi_K$. Specifically, if ψ is not onto, then φ cannot be onto. For clarity, we also give a direct proof. For each $\ell \in L$, we must find $Kg \in G/K$ such that $\psi(Kg) = \ell$. Since $L = \varphi(G)$, it follows that there exists $g \in G$ such that $\varphi(g) = \ell$. Now, by definition of φ_K, we know that $\varphi_K(g) = Kg$. Since $\varphi = \psi \circ \varphi_K$, we see that

$$\varphi(g) = \psi(\varphi_K(g)) = \psi(Kg).$$

Hence $\psi(Kg) = \ell$ and so ψ is onto.

Second, we verify that ψ is one-to-one. We will use the following lemma in the proof.

Lemma 4.13. *Let $\varphi\colon G \to H$ be a homomorphism of groups. Then the following are equivalent:*

(a) *φ is one-to-one.*

(b) $\ker(\varphi) = \{1\}$.

The reader should compare this lemma with a well known lemma in linear algebra.

Lemma 4.14. *Let $L\colon V \to W$ be a linear function between vector spaces V,W. Then the following are equivalent:*

(a) *L is one-to-one.*

(b) $\ker(L) = \{0\}$.

We will now prove that ψ is one-to-one. It is enough to prove that $\ker(\psi) = \{1\}$. Given $Kg \in \ker(\psi)$, by definition, we have $\psi(Kg) = 1$. Since $\varphi = \psi \circ \varphi_K$, we see that $\varphi(g) = 1$ and so $g \in K = \ker(\varphi)$. However, if $g \in K$, then $Kg = K$. Hence, $\ker(\psi) = \{1\}$; remember that the trivial coset K is the identity element in G/K.

We now summarize the above:

Theorem 4.15 (First Isomorphism Theorem)**.** *Let G,H be groups, $\varphi\colon G \to H$ a homomorphism, $L = \varphi(G)$, $K = \ker(\varphi)$, and $\varphi_K\colon G \to G/K$ the quotient homomorphism. Then there exists a bijective group homomorphism $\psi\colon G/K \to L$ such that $\varphi = \psi \circ \varphi_K$. In particular, $G/K \cong L$. This can be represented by the commutative diagram:*

$$\begin{tikzcd}
G \arrow[rr, "\varphi"] \arrow[dr, "\varphi_K"'] & & L = \varphi(H) \leq H \\
& G/K \arrow[ur, dashed, "\psi"'] &
\end{tikzcd}$$

To appreciate the first isomorphism theorem, one might consider thinking of it as a type of change of variables result. In a change of variable operation (e.g. a u–substitution in integration theory), we make a change of variable to simplify our problem. Gauss–Jordan elimination is an example of such a result for systems of linear equations. In the context of group theory, the group G/K is defined in an abstract way that is internal to G. It can be difficult to understand the group G/K as a result. The subgroup L of H might be more concrete as the group H is external to G in some sense. Because G/K and L are isomorphic, from the view of just group theory, G/K and L are the same group but are two different realizations of this group. In particular, if our interest was in the group G/K, we could instead study the group L. Any property about groups that is satisfied by L will be satisfied by G/K.

For instance, we can define a function $\varphi\colon \mathbf{Z} \to \{\pm 1\}$. We will view $\{1,-1\}$ as a group under multiplication (e.g. just regular multiplication by numbers). We could take $\{\pm \mathrm{I}_m\}$ for any $m \in \mathbf{N}$ where $\mathrm{I}_m \in \mathrm{GL}(m,\mathbf{R})$ is the m by m identity matrix. These are all different concrete realizations of the same group (e.g. all of these groups are isomorphic). Regardless of which realizations we pick, we have the function

$$\varphi\colon \mathbf{Z} \to \{\pm 1\}, \quad \varphi(n) = \begin{cases} 1, & n \text{ is even}, \\ -1, & n \text{ is odd}. \end{cases}$$

It is a simple matter to verify that φ is a group homomorphism and $\ker(\varphi) = 2\mathbf{Z}$, the subgroup of even numbers. By the first isomorphism theorem, we see that $\mathbf{Z}/2\mathbf{Z} \cong \{\pm 1\}$.

If $A_m \in \mathrm{SL}(2,\mathbf{R})$ is given by

$$A_m = \begin{pmatrix} \cos(2\pi/m) & \sin(2\pi/m) \\ -\sin(2\pi/m) & \cos(2\pi/m) \end{pmatrix},$$

we see that $A_m^m = \mathrm{I}_2$ and $A_m^j \neq A_m^i$ for all $1 \leq i,j \leq m-1$ with $i \neq j$. In particular, the set

$$G_m = \left\{ \mathrm{I}_2 = A_m^0, A_m, A_m^2, \dots, A_m^{m-2}, A_m^{m-1} \right\}$$

is a subgroup of $\mathrm{SL}(2,\mathbf{R})$ and $|G_m| = m$. We can define a surjective homomorphism $\varphi_m\colon \mathbf{Z} \to G_m$ by

$$\varphi_m(n) = A_m^i, \quad n = qm + i, \quad 0 \leq i \leq m-1.$$

The kernel of φ_m is $m\mathbf{Z}$ and so $G_m \cong \mathbf{Z}/m\mathbf{Z}$ by the first isomorphism theorem. In particular, we can think of the quotient group $\mathbf{Z}/m\mathbf{Z}$, which is formally a set of cosets, as the group of rotations generated by a rotation of order m. We can also view $\mathbf{Z}/m\mathbf{Z}$ via modular arithmetic, which in its full scope uses the ring structure of $\mathbf{Z}$ and not merely the group structure. We will cover this example more extensively when studying ideals in rings which plays the analog of a normal subgroup of a group.

The following corollary of the First Isomorphism Theorem is the analog of the fact that given any surjective linear function $L\colon V \to W$, there exists a basis $\mathscr{B}_V = \{v_1, \dots v_m\}$ of V and a basis $\mathscr{B}_W = \{w_1, \dots, w_m\}$ such that $L(v_i) = w_i$.

Corollary 4.16. *Let G, H be groups, $\psi\colon G \to H$ a surjective group homomorphism, and $K = \ker\psi$. Then H and G/K are isomorphic and the diagram*

$$\begin{tikzcd}
G \arrow[rr, bend left, "\psi"] \arrow[dr, "\psi_K"'] & & H \\
& G/K \arrow[ur, "\overline{\psi}"'] &
\end{tikzcd}$$

commutes. Namely, $\psi = \overline{\psi} \circ \psi_K$.

4.4.2 The Second Isomorphism Theorem

The following result is often referred to as the **Second Isomorphism Theorem**.

Theorem 4.17 (Second Isomorphism Theorem)**.** *Let G be a group, $K \leq G$, and $H \lhd G$. Then*

(a) The set
$$HK = \{hk \ : \ h \in H, \ k \in K\}$$
is a subgroup of G.

(b) $H \cap K$ is a normal subgroup of K.

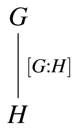
(c) The groups HK/H and $K/(H \cap K)$ are isomorphic.

Proof. For (a), since both H,K are subgroups of G, we know that $e \in H$ and $e \in K$. In particular, $e \in HK$. Given $h_1k_1, h_2k_2 \in HK$, we see that
$$h_1k_1(h_2k_2)^{-1} = h_1k_1k_2^{-1}h_2^{-1}.$$
Since H is normal, we know that $(k_1k_2^{-1})H = H(k_1k_2^{-1})$. Hence, there exists $h_3 \in H$ such that $k_1k_2^{-1}h_2 = h_3k_1k_2^{-1}$. Therefore,
$$h_1k_1k_2^{-1}h_2^{-1} = h_1h_3k_1k_2^{-1}.$$
Since H,K are subgroups, $h_1h_3 \in H$ and $k_1k_2^{-1} \in K$. In particular, $h_1h_3k_1k_2^{-1} \in HK$ and so $h_1k_1(h_2k_2)^{-1} \in HK$. It now follows that HK is a subgroup.

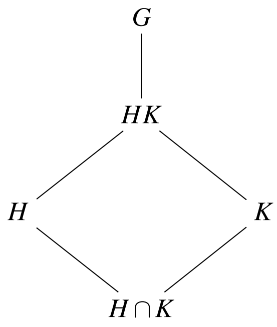
For (b), we know that $H \cap K$ is a subgroup of G and so we need only verify that it is normal. For that, given $h \in H \cap K$ and $k \in K$, we must prove that $k^{-1}hk \in H \cap K$. Since H is normal in G and $h \in H$, it follows that $k^{-1}hk \in H$. Since K is a subgroup of G and $h,k \in K$, it follows that $k^{-1}hk \in K$. Thus, $k^{-1}hk \in H \cap K$.

For (c), we will construct an isomorphism $\psi \colon K/(H \cap K) \to HK/H$. By definition of HK, $K \leq HK$. Taking $\psi_H \colon HK \to HK/H$ to be the canonical homomorphism, the restriction of ψ_H to K is a homomorphism. Since $\ker \psi_H = H$, the kernel of the restriction of ψ_H to K is $H \cap K$. By the First Isomorphism Theorem, it follows that $\psi_H(K)$ is isomorphic to $K/(H \cap K)$. It remains to show that the restriction of ψ_H to K is surjective. For that, given $kH \in HK/H$, we must find $k_1 \in K$ such that $\psi_H(k_1) = kH$. First, we can write $k = h_1k_1$ for $h_1 \in H$ and $k_1 \in K$. Since H is normal, we know that $k_1H = Hk_1$ and so $h_1k_1 = k_1h_2$ for some $h_2 \in H$. In particular, we see that $k = k_1h_2$ and so $kk_1^{-1} \in H$. Hence, $kH = k_1H$. By definition of ψ_H, we have $\psi_H(k_1) = k_1H = kH$. Therefore, $\psi_H(K) = HK/H$ and so $K/(H \cap K)$ and HK/H are isomorphic. □

When $H \leq G$, it is common to associate to G,H a diagram



We have the following "diamond" associated to the Second Isomorphism Theorem:



By (c) in Theorem 4.17 the opposite sides of the diamond

$$\begin{tikzcd}
 & G \arrow[d, dash] & \\
 & HK \arrow[dl, equal] \arrow[dr, dash] & \\
H \arrow[dr, dash] & & K \arrow[dl, equal] \\
 & H \cap K &
\end{tikzcd}$$

are the “same”. Theorem 4.17 is also sometimes called the diamond isomorphism theorem.

4.4.3 The Third Isomorphism Theorem

The following result is often referred to as the **Third Isomorphism Theorem**. It is essentially a conglomerate of observations about the subgroup group structure of G and the subgroup structure of the quotient group G/H for $H \lhd G$. The most note worth of these results is (e).

Theorem 4.18 (Third Isomorphism Theorem)**.** *Let G be a group and $H \lhd G$.*

(a) If $K \leq G$ and $H \subseteq K \subseteq G$, then K/H is a subgroup of G/H.

(b) Every subgroup of G/H is of the form K/H, for some $K \leq G$ such that $H \subseteq K \subseteq G$.

(c) If $K \lhd G$ and $H \subseteq K \subseteq G$, then K/H is a normal subgroup of G/H.

(d) Every normal subgroup of G/H is of the form K/H, for some $K \lhd G$ such that $H \subseteq K \subseteq G$.

(e) If $K \lhd G$ and $H \subseteq K \subseteq G$, then the groups $(G/H)/(K/H)$ and G/K are isomorphic.

Proof. For (a), we can restrict the canonical homomorphism $\psi_H \colon G \to G/H$ to K. The image $\psi_H(K) = K/H$ and $\psi_H(K) \leq G/H$.

For (b), given a subgroup $L \leq G/H$, we assert that the pullback $\psi_H^{-1}(L)$ of L is a subgroup of G. Recall,

$$\psi_H^{-1}(L) \overset{\text{def}}{=} \{g \in G \;:\; \psi_H(g) \in L\}.$$

Since L is s subgroup, $e_{G/H} \in L$. As $\psi_H(e_G) = e_{G/H}$, we see that $e_G \in \psi_H^{-1}(L)$. Given $g_1, g_2 \in \psi_H^{-1}(L)$, there exist $\ell_1, \ell_2 \in L$ such that $\psi_H(g_1) = \ell_1$ and $\psi_H(g_2) = \ell_2$. Since L is a subgroup of G/H, we have $\ell_1 \ell_2^{-1} \in L$. Additionally, we have

$$\psi_H(g_1 g_2^{-1}) = \psi_H(g_1)(\psi_H(g_2))^{-1} = \ell_1 \ell_2^{-1} \in L.$$

Hence $g_1 g_2^{-1} \in \psi_H^{-1}(L)$ and so $\psi_H^{-1}(L)$ is a subgroup of G. By definition, $\psi_H(\psi_H^{-1}(L)) = L = \psi_H^{-1}(L)/H$, as needed to verify (b).

For (c), given $kH \in K/H$ and $g \in G/H$, we must show that $(gH)^{-1}kHgH \in K/H$. Since K is normal, we know that $g^{-1}kg = k_1 \in K$. As $\psi_H(k) = kH$ and $\psi_H(g) = gH$, we see that

$$(gH)^{-1}kHgH = (\psi_H(g))^{-1}\psi_H(k)\psi_H(g) = \psi_H(g^{-1}kg) = \psi_H(k_1) \in K/H.$$

For (d), given $L \lhd G/H$, we assert that $\psi_H^{-1}(L) \lhd G$. Given $g_0 \in \psi_H^{-1}(L)$ and $g \in G$, we must prove that $g^{-1}g_0 g \in \psi_H^{-1}(L)$. First, since $g_0 \in \psi_H^{-1}(L)$, there exists $\ell_0 \in L$ such that $\psi_H(g_0) = \ell_0$. Now, we have

$$\psi_H(g^{-1}g_0 g) = (\psi_H(g))^{-1}\ell_0(\psi_H(g)) \in L$$

since L is normal. Hence, $g^{-1}g_0g \in \psi_H^{-1}(L)$. As in (b), we have $\psi_H(\psi_H^{-1}(L)) = L = \psi_H^{-1}(L)/H$.

For (e), we have

$$\begin{tikzcd}
G \arrow[rr, bend left, "\psi"] \arrow[r, bend right, "\psi_H"'] & G/H \arrow[r, bend right, "\psi_{K/H}"'] & (G/H)/(K/H)
\end{tikzcd}$$

where $\psi = \psi_{K/H} \circ \psi_H$. By the First Isomorphism Theorem, we know that $\psi(G)$ and $G/\ker\psi$ are isomorphic. Since both ψ_H and $\psi_{K/H}$ are surjective, it follows that ψ is surjective. In particular, $\psi(G) = (G/H)/(K/H)$. Given $g \in \ker\psi$, since $\ker\psi_{K/H} = K/H$, we must have $\psi_H(g) = \ker\psi_{K/H} = K/H$. Therefore, $g \in \psi_H^{-1}(K/H) = K$. Hence, $\ker\psi = K$, as needed. □

It seems to be popular to point out one aesthetically appealing (notationally) view of (e) as an analog of fractional cancellation. Specifically, if we write $G/H = \frac{G}{H}$, then (e) of Theorem 4.18 asserts that

$$\frac{\frac{G}{H}}{\frac{K}{H}} \cong \frac{G}{K}.$$

5 Groups: Examples

For concreteness sake, we will first only consider concrete, finite dimension, real vector spaces. Specifically, we will work with $\mathbf{R}^n$ viewed as the n–fold product of $\mathbf{R}$.

5.1 Basics of $\mathbf{R}^n$ and linear functions

We define $\mathbf{R}^n$ to be the set

$$\mathbf{R}^n = \{(x_1, \dots, x_n) \ : \ x_1, \dots, x_n \in \mathbf{R}\}.$$

We call $x \in \mathbf{R}^n$ a **n–vector** and we can view $x = (x_1, \dots, x_n)$. The **zero vector** is $0 = (0, 0, \dots, 0)$. We can add and scalar multiply vectors. Given $x, y \in \mathbf{R}^n$ with $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, we define

$$x + y = (x_1 + y_1, \dots, x_n + y_n) \in \mathbf{R}^n.$$

Given $\alpha \in \mathbf{R}$ and $x \in \mathbf{R}^n$ with $x = (x_1, \dots, x_n)$, we define

$$\alpha x = (\alpha x_1, \dots, \alpha x_n) \in \mathbf{R}^n.$$

A function $L\colon \mathbf{R}^m \to \mathbf{R}^n$ is **linear** if for each $x, y \in \mathbf{R}^n$ and $\alpha, \beta \in \mathbf{R}^n$, we have

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y).$$

We define $\mathrm{Hom}(\mathbf{R}^m, \mathbf{R}^n)$ to be to set of all linear functions $L\colon \mathbf{R}^m \to \mathbf{R}^n$. This set is a real vector space via the addition and scalar operations

$$(L_1 + L_2)(x) = L_1(x) + L_2(x)$$

and

$$(\alpha L)(x) = \alpha L(x)$$

for $L, L_1, L_2 \in \mathrm{Hom}(\mathbf{R}^m, \mathbf{R}^n)$ and $\alpha \in \mathbf{R}$.

Taking $e_i = (0, 0, \dots, 0, 1, 0, \dots, 0)$ where the 1 occurs at the ith coordinate, every vector $x \in \mathbf{R}^n$ can be expressed as

$$x = \sum_{i=1}^{n} x_i e_i.$$

5.2 Basics: Matrices and Linear Functions

Recall that an n by m real matrix is an array with n rows and m columns of real numbers. That is, something of the form

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,m} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n_3} & \dots & a_{n,m} \end{pmatrix}.$$

We will use $A = (a_{i,j})$ as shorthand notation for the full matrix expression above. One remembers that for $a_{i,j}$ that i parameterizes the rows of A and j parameterizes the columns of A. We can think of this array/matrix A as either m vectors in $\mathbf{R}^n$ (via the columns) or n vectors in $\mathbf{R}^m$ (via the rows). We define the column vectors $v_1, \dots, v_m \in \mathbf{R}^n$ of the matrix A to be

$$v_i = (a_{1,i}, \dots, a_{n,i}) \in \mathbf{R}^n$$

and the row vectors $v^1,\dots,v^n \in \mathbf{R}^m$ of the matrix A to be

$$v^i = (a_{i,1},\dots,a_{i,m}) \in \mathbf{R}^m.$$

From this matrix, we can define two linear functions from A using either the columns or rows. However, before doing constructing these linear functions, we briefly review the theory of linear functions $L\colon \mathbf{R}^m \to \mathbf{R}^n$.

Given a linear function $L\colon \mathbf{R}^m \to \mathbf{R}^n$ and a vector $x \in \mathbf{R}^m$, we know that

$$x = \sum_{i=1}^m x_i e_i$$

and so

$$L(x) = L\left(\sum_{i=1}^m x_i e_i\right) = \sum_{i=1}^m x_i L(e_i).$$

Hence, the value of x under L is completely determined by the values of $e_1,\dots,e_m$ under L. Moreover, given vectors $v_1,\dots,v_m \in \mathbf{R}^m$, we can construct a linear function $L\colon \mathbf{R}^m \to \mathbf{R}^n$ by

$$L(x) = \sum_{i=1}^m x_i v_i.$$

Under this construction, we see that

$$L(e_i) = v_i.$$

Hence, the set of linear functions $L\colon \mathbf{R}^m \to \mathbf{R}^n$ is the same as the set of choices of m vectors in $\mathbf{R}^n$. We can view this choice of m vectors in $\mathbf{R}^n$ as an n by m real matrix by taking the columns of this array to be the m vectors viewed as column vectors.

We can view a linear function $L\colon \mathbf{R}^n \to \mathbf{R}^m$ as n choices of vectors in $\mathbf{R}^m$. Again, every vector $x \in \mathbf{R}^n$ can be written as

$$x = \sum_{i=1}^n x_i e_i,$$

and so we have

$$L(x) = L\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i L(e_i).$$

For any selection of vectors $v^1,\dots,v^n \in \mathbf{R}^m$, we have a linear function $L^T : \mathbf{R}^n \to \mathbf{R}^m$ defined by

$$L^T(x) = \sum_{i=1}^n x_i v^i.$$

We can form an n by m matrix by taking the vectors $v^1,\dots,v^n$ to be the rows of the n by m array.

Summary: Given an n by m real matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,m} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n_3} & \dots & a_{n,m} \end{pmatrix}.$$

with column vectors $v_1,\dots,v_m$ defined by

$$v_i = (a_{1,i},\dots,a_{n,i}) \in \mathbf{R}^n$$

and row vectors $v^1,\dots,v^n$ defined by

$$v^i = (a_{i,1},\dots,a_{i,m}) \in \mathbf{R}^m,$$

we can view A as a linear function $L\colon \mathbf{R}^m \to \mathbf{R}^n$ defined by $L(e_i) = v_i$ or as a linear function $L^T\colon \mathbf{R}^n \to \mathbf{R}^m$ defined by $L^T(e_i) = v^i$. The linear function L^T is called the **dual map of** L or **adjoint of** L.

If we denote by $\mathrm{M}(n,m;\mathbf{R})$ the set of n by m real matrices, we see now that there is a bijection between $\mathrm{Hom}(\mathbf{R}^m,\mathbf{R}^n)$ and $\mathrm{M}(n,m;\mathbf{R})$. The bijection is given by sending the linear function $L\colon \mathbf{R}^m \to \mathbf{R}^n$ to the matrix A with column vectors v_i given by $L(e_i)$. We also have a bijection between $\mathrm{Hom}(\mathbf{R}^n,\mathbf{R}^m)$ and $\mathrm{M}(n,m;\mathbf{R})$ given by sending the linear function $L\colon \mathbf{R}^n \to \mathbf{R}^m$ to the matrix with row vectors $v^i = L(e_i)$.

This relationship between $\mathrm{Hom}(\mathbf{R}^m,\mathbf{R}^n)$ and $\mathrm{Hom}(\mathbf{R}^n,\mathbf{R}^m)$ can also be connected through a relationship between $\mathrm{M}(n,m;\mathbf{R})$ and $\mathrm{M}(m,n;\mathbf{R})$. Given a matrix $A \in \mathrm{M}(n,m;\mathbf{R})$ with $A = (a_{i,j})$, the **transpose of** A is an m by n matrix $A^T \in \mathrm{M}(m,n;\mathbf{R})$ defined by $A^T = (a_{j,i})$. Concretely, we form a matrix A^T by using the column vectors v_i of A as the row vectors of A^T. Alternatively, we form A^T by using the row vectors v^i of A as the column vectors of A^T. The transpose operation gives a bijections between the set of n by m matrices $\mathrm{M}(n,m;\mathbf{R})$ and the set of m by n matrices $\mathrm{M}(m,n;\mathbf{R})$.

The sets $\mathrm{M}(n,m;\mathbf{R})$ and $\mathrm{Hom}(\mathbf{R}^m,\mathbf{R}^n)$ be has natural vector space structures.

5.2.1 Vector addition and scalar multiplication for $\mathrm{M}(n,m;\mathbf{R})$.

Given $A,B \in \mathrm{M}(n,m;\mathbf{R})$, we write $A = (a_{i,j})$ and $B = (b_{i,j})$ where $a_{i,j}$ is the coefficient in the ith row and jth column of A. We define

$$A + B = C = (c_{i,j}), \qquad \text{where } c_{i,j} = a_{i,j} + b_{i,j}$$

and for $\alpha \in \mathbf{R}$, we define

$$\alpha A = (\alpha a_{i,j}).$$

The vector vector is played by $0_{n,m}$ where the (i,j)–coefficient is 0 for all i,j.

5.2.2 Vector addition and scalar multiplication for $\mathrm{Hom}(\mathbf{R}^m,\mathbf{R}^n)$.

Given $L_1, L_2 \in \mathrm{Hom}(\mathbf{R}^m,\mathbf{R}^n)$, we define vector addition by

$$(L_1 + L_2)(x) = L_1(x) + L_2(x)$$

and for $\alpha \in \mathbf{R}$ and $L \in \mathrm{Hom}(\mathbf{R}^m,\mathbf{R}^n)$, we define scalar multiplication by

$$(\alpha L)(x) = \alpha L(x).$$

5.2.3 Dot product or Euclidean inner product

Given vectors $x,y \in \mathbf{R}^m$ where $x = (x_1,\dots,x_m)$ and $y = (y_1,\dots,y_m)$, we define the inner product or **dot product** of x,y by

$$x \cdot y = x_1y_1 + x_2y_2 + \cdots + x_my_m = \sum_{i=1}^m x_iy_i.$$

5.2.4 Matrix muliplication

Here, we consider matrix multiplication. Given $A \in \mathrm{M}(n,m;\mathbf{R})$ and $B \in \mathrm{M}(m,p;\mathbf{R})$, we define

$$C = AB = (c_{i,j})$$

where

$$c_{i,j} = a^i \cdot b_j = \sum_{k=1}^{m} a_{k,i} b_{j,k}$$

where a^i is the ith row of A and b_j is the jth column of B. Note that since A is an n by m matrix, the matrix A has n row vectors $a^1, \dots, a^n \in \mathbf{R}^m$ and B being an m by p matrix has p column vectors $b_1, \dots, b_p \in \mathbf{R}^m$. In particular, since $a^i, b_j \in \mathbf{R}^m$, then $a^i \cdot b_j$ is well defined. We see that $C \in \mathrm{M}(n,p;\mathbf{R})$.

5.2.5 Square matrices

When $m = n$, we write $\mathrm{M}(m,\mathbf{R})$ for the set of m by m matrices. Given $A, B \in \mathrm{M}(m,\mathbf{R})$, we see that $AB \in \mathrm{M}(m,\mathbf{R})$. In particular, $\mathrm{M}(m,\mathbf{R})$ has a binary operator given by matrix multiplication. One can check that it is associative and has an identity given by the m by m identity matrix I_m. This can be defined as follows. First, define

$$\delta_{i,j} = \begin{cases} 0, & \text{if } i \neq j, \\ 1, & \text{if } i = j. \end{cases}.$$

This function is sometimes called the **Kronecker δ–function**. We define

$$\mathrm{I}_m = (\delta_{i,j}).$$

The **determinant** of a matrix $A \in \mathrm{M}(m,\mathbf{R})$ can be defined recursively from the base case $m = 2$. The determinant can be viewed as a function $\det\colon \mathrm{M}(m;\mathbf{R}) \to \mathbf{R}$. When $m = 2$, we have

$$\det(A) = ad - bc, \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Proposition 5.1. *$A \in \mathrm{M}(m,\mathbf{R})$ has a multiplicative inverse if and only if $\det(A) \neq 0$.*

Both $\mathrm{M}(m,\mathbf{R})$ and $\mathbf{R}$ are monoids under multiplication.

Lemma 5.2. *If $A, B \in \mathrm{M}(m,\mathbf{R})$, then*

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

Also, $\det(\mathrm{I}_m) = 1$.

As a result of this lemma, we see that det is a homomorphism of monoids.

5.3 The general and special linear groups

We define $\mathrm{GL}(m,\mathbf{R})$ to be the group of m by m real matrices with non-zero determinant. That is

$$\mathrm{GL}(m,\mathbf{R}) = \{A \in \mathrm{M}(m,\mathbf{R}) \ : \ \det(A) \neq 0\}.$$

Since A has an inverse $A^{-1} \in \mathrm{M}(m,\mathbf{R})$, we see that if $A \in \mathrm{GL}(m,\mathbf{R})$ then $A^{-1} \in \mathrm{GL}(m,\mathbf{R})$. To see this claim, note that $AA^{-1} = \mathrm{I}_m$ and so

$$\det(AA^{-1}) = \det(A)\det(A^{-1}) = \det(\mathrm{I}_m) = 1.$$

Since $\det(A) \neq 0$, we can divide both sides of the equation

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

by $\det(A)$, and obtain

$$\det(A^{-1}) = \frac{1}{\det(A)} = (\det(A))^{-1} \neq 0.$$

Since matrix multiplication is associative, we see that $\mathrm{GL}(m,\mathbf{R})$ is a group under multiplication since we have an identity and inverses. The group $\mathrm{GL}(m,\mathbf{R})$ is called the **general linear group**.

The subset $\mathrm{SL}(m,\mathbf{R})$ of $\mathrm{GL}(m,\mathbf{R})$ given by

$$\mathrm{SL}(m,\mathbf{R}) = \{A \in \mathrm{GL}(m,\mathbf{R}) \ : \ \det(A) = 1\}$$

is a subgroup since $\mathrm{I}_m \in \mathrm{SL}(m,\mathbf{R})$, since $\det(A^{-1}) = 1$ when $\det(A) = 1$, and since $\det(AB) = 1$ when $\det(A) = \det(B) = 1$. The subgroup $\mathrm{SL}(m,\mathbf{R})$ is called the **special linear group**.

We can view $\det\colon \mathrm{GL}(m,\mathbf{R}) \to \mathbf{R}^\times = \mathbf{R} - \{0\}$ as a homomorphism of groups where $\mathbf{R}^\times$ is a group under multiplication with identity 1. The kernel of det is $\mathrm{SL}(m,\mathbf{R})$ since

$$\ker(\det) = \{A \in \mathrm{GL}(m,\mathbf{R}) \ : \ \det(A) = 1\}.$$

5.3.1 Special types of matrices

Definition 61 (Diagonal matrix)**.** Given a matrix $A \in \mathrm{M}(m,\mathbf{R})$ with $A = (a_{i,j})$, we say that A is **diagonal** if $a_{i,j} = 0$ for all $i \neq j$.

Definition 62 (Upper/Lower triangular matrix)**.** Given a matrix $A \in \mathrm{M}(m,\mathbf{R})$ with $A = (a_{i,j})$, we say that A is **upper triangular** if $a_{i,j} = 0$ for all $i \geq j$. We say that A is lower triangular if $a_{i,j} = 0$ for all $i \leq j$.

Definition 63 (Unipotent matrix)**.** Given an upper triangular matrix $A \in \mathrm{M}(m,\mathbf{R})$ with $A = (a_{i,j})$, we say that A is **upper unipotent** if $a_{i,i} = 1$ for all i. We say that A is lower unipotent if $a_{i,i} = 1$ for all i.

Definition 64 (Orthogonal matrix)**.** Given a matrix $A \in \mathrm{M}(m,\mathbf{R})$, we say that A is an **orthgonal matrix** if for each $x, y \in \mathbf{R}^m$, we have

$$(Ax)\cdot(Ay) = x \cdot y.$$

Definition 65 (Transpose)**.** Given a matrix $A \in \mathrm{M}(n,m;\mathbf{R})$, the **transpose of A** is a matrix $A^T \in \mathrm{M}(m,n;\mathbf{R})$ where $A^T = (b_{i,j})$ is given by

$$b_{i,j} = a_{j,i}.$$

When $A \in \mathrm{M}(m,\mathbf{R})$, then $A^T \in \mathrm{M}(m,\mathbf{R})$. We also have the following facts.

Lemma 5.3. *If $A \in \mathrm{M}(n,m;\mathbf{R})$, then $(A^T)^T = A$.*

Lemma 5.4. *If $A, B \in \mathrm{M}(m,\mathbf{R})$, then $(AB)^T = B^T A^T$. If $A \in \mathrm{GL}(n,\mathbf{R})$, then $(A^{-1})^T = (A^T)^{-1}$. In particular, if $A \in \mathrm{GL}(m,\mathbf{R})$, then $A^T \in \mathrm{GL}(m,\mathbf{R})$. In fact, $\det(A) = \det(A^T)$.*

Definition 66 (Symmetric/Skew-Symmetric matrix)**.** Given a matrix $A \in \mathrm{M}(m,\mathbf{R})$, we say that A is **symmetric** if $A^T = A$. We say that A is **skew symmetric** if $A^T = -A$.

Lemma 5.5. *Every matrix $A \in \mathrm{M}(m,\mathbf{R})$ can be written as*

$$A = A_s + A_{ss}$$

where A_s is a symmetric matrix and A_{ss} is a skew symmetric matrix.

Recall that a collection of vectors $v_1, \dots, v_k \in \mathbf{R}^m$, we say $\{v_1, \dots, v_k\}$ is an **orthogonal set** if $v_i \cdot v_j = 0$ for all $i \neq j$. We say that set $\{v_1, \dots, v_k\}$ is an **orthonormal set** if $v_i \cdot v_j = \delta_{i,j}$ where $\delta_{i,j}$ is the Kronecker δ–function.

Proposition 5.6. *Given $A \in \mathrm{M}(m,\mathbf{R})$, then following are equivalent:*

(i) A is orthogonal.

(ii) $AA^T = \mathrm{I}_m$ (or equivalently, $A^T = A^{-1}$).

(iii) The columns of A, $\{a_1, \dots, a_m\}$ are an orthonormal set.

(iv) The rows of A, $\{a^1, \dots, a^m\}$ are an orthonormal set.

5.3.2 $\mathrm{O}(2)$ and $\mathrm{SO}(2)$

There are lots of different subgroups of $\mathrm{GL}(m,\mathbf{R})$. We have seen one already, namely $\mathrm{SL}(m,\mathbf{R})$. Before giving some general families, we will first look at a few concrete examples when $m=2$. One example has a connection Euclidean geometry. We define

$$\mathrm{SO}(2)=\left\{R_\theta=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix} \;:\; \theta\in[0,2\pi)\right\}.$$

This group is called the 2–dimensional special orthogonal group or **circle group**. We can view the elements R_θ of $\mathrm{SO}(2)$ as simply a specific choice of the parameter θ. When $\theta=\pi/2$, we get the matrix

$$R_{\pi/2}=\begin{pmatrix}\cos(\pi/2) & \sin(\pi/2)\\ -\sin(\pi/2) & \cos(\pi/2)\end{pmatrix}=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}.$$

We see that $R_{\pi/2}(e_1)=-e_2$ and $R_{\pi/2}(e_2)=e_1$. This is a clockwise **rotation** about the origin in $\mathbf{R}^2$ by $\pi/2$. We then might guess that $R_{\pi/2}R_{\pi/2}=R_\pi$. First, we have

$$\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}=\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$$

A rotation by π would send $e_1\to -e_1$ and send $e_2\to -e_2$. We should also expect $R_\pi R_\pi=R_{2\pi}=R_0$. In this case we have

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

A rotation by 0 is equivalent to a rotation by 2π. This would be the identity, which is what we got.

Summary: The group $\mathrm{SO}(2)=\{R_\theta \;:\; \theta\in[0,2\pi)\}$ is the group of rotations in $\mathbf{R}^2$. We have a bijection between the unit circle, which we denote by C in $\mathbf{R}^2$ and the group $\mathrm{SO}(2)$. The association $\mathrm{SO}(2)\to C$ is given by $R_\theta\to R_\theta(e_1)$. Viewing e_1 as the point $(1,0)\in\mathbf{R}^2$, this point lies on the unit circle C. Every rotation R_θ takes points on the unit circle to points on the unit circle. It should be clear that a rotation is completely determined by where it sends the point $e_1=(1,0)$. We see also that $R_\theta(e_1)=(\cos\theta,-\sin\theta)$. That $(\cos\theta,-\sin\theta)$ lies on C is seen by verifying the x,y coefficients are solutions to the equation $x^2+y^2=1$. In this case, we have

$$\cos^2\theta+\sin^2\theta=1,$$

which holds either by definition of $\cos,\sin$ or the pythagorean identity. Note, via the same logic, we also have

$$\det(R_\theta)=\cos^2\theta+\sin^2\theta=1.$$

One can check that $R_\theta(R_\theta)^T=\mathrm{I}_2$. This requires one verify

$$\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

The reader can (privately) verify this; it does not require any heavy lifting. In particular, R_θ is an orthogonal matrix.

5.3.3 A geometric view of orthogonal matrices

One might wonder if every orthogonal matrix in $\mathrm{M}(2,\mathbf{R})$ is an element of $\mathrm{SO}(2)$. For that, we define

$$\mathrm{O}(2)=\{A\in\mathrm{M}(2,\mathbf{R}) \;:\; A \text{ is orthogonal}\}.$$

From above, we see that $\mathrm{SO}(2) \subset \mathrm{O}(2)$ and the question is whether or not $\mathrm{O}(2) = \mathrm{SO}(2)$. Consider

$$\tau = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Since the columns of τ form an orthonormal set, we see that $\tau \in \mathrm{O}(2)$. However, we see that $\det(\tau) = -1$. As all of the elements of $\mathrm{SO}(2)$ have determinant 1, we see that $\tau \notin \mathrm{SO}(2)$.

Even though $\mathrm{O}(2)$ is bigger than $\mathrm{SO}(2)$, our hope that $\mathrm{O}(2) = \mathrm{SO}(2)$ was only slightly wrong. We define

$$\tau\,\mathrm{SO}(2) = \{\tau A \ : \ A \in \mathrm{SO}(2)\}.$$

This is merely the set of matrices of the form τA where $A \in \mathrm{SO}(2)$. Note that if $B \in \tau\,\mathrm{SO}(2)$, then $B = \tau A$ for some $A \in \mathrm{SO}(2)$. Consequently, $\det(B) = -1$ since

$$\det(B) = \det(\tau A) = \det(\tau)\det(A) = -1.$$

Remembering that $\mathrm{SO}(2)$ and $\tau\,\mathrm{SO}(2)$ are just subsets of $\mathrm{M}(2, \mathbf{R})$, I claim that

$$\mathrm{SO}(2) \cap \tau\,\mathrm{SO}(2) = \emptyset.$$

To prove this claim, we argue as follows. Let $B \in \mathrm{SO}(2) \cap \tau\,\mathrm{SO}(2)$. First, we have $B \in \mathrm{SO}(2)$, and so $\det(B) = 1$. Second, we also have $B \in \tau\,\mathrm{SO}(2)$ and so $\det(B) = -1$. As $1 \neq -1$, we see that no such matrix B could exist and so

$$\mathrm{SO}(2) \cap \tau\,\mathrm{SO}(2) = \emptyset.$$

Proposition 5.7.

$$\mathrm{O}(2) = \mathrm{SO}(2) \cup \tau\,\mathrm{SO}(2).$$

In particular, if A is an orthogonal matrix and $\det(A) = 1$*, then $A = R_\theta$ for some $\theta \in [0, 2\pi)$. If $A \in \mathrm{M}(2, \mathbf{R})$ is orthogonal and* $\det(A) = -1$*, then $A = \tau R_\theta$ for some $\theta \in [0, 2\pi)$.*

The matrix τ represents the reflection of $\mathbf{R}^2$ across the y–axis. Explicitly, the linear function is $L(x,y) = (-x,y)$. We can restate the above proposition more plainly.

Proposition 5.8. *Every element in $\mathrm{O}(2)$ is represented by either a rotation or a rotation followed by reflection across the y–axis.*

We now return to further investigate what exactly does an orthogonal matrix represent. We can view the matrix as a linear function $L \colon \mathbf{R}^2 \to \mathbf{R}^2$. The condition for L to be orthogonal is that given any two vectors $x, y \in \mathbf{R}^2$, we have $L(x) \cdot L(y) = x \cdot y$. The angle between two vectors x, y is completely determined by $x \cdot x$, $y \cdot y$, and $x \cdot y$. The length of a vector x is given by $\sqrt{x \cdot x}$. Now, our function L sends each pair of vectors x, y in $\mathbf{R}^2$ to a pair of vectors $L(x), L(y)$ such that $x, L(x)$ have the same length, $y, L(y)$ have the same length, and the angle between x, y is the same as the angle between $L(x), L(y)$. This is a lot to ask for. Rotations about the origin are very concrete examples of linear functions. They are also examples of functions the preserve lengths and angles. Reflections across any line that pass through the origin are also examples of linear functions that preserve angles and norms. Notice that given any line through the origin, there is a unique rotation that takes the y–axis to this line. In particular, we can think of $\tau\,\mathrm{SO}(2)$ as the set of reflections across any line through the origin. So we can restate the proposition once again.

Proposition 5.9. *Every $A \in \mathrm{O}(2)$ is either a rotation about the origin or the reflection across some line through the origin.*

5.3.4 The translations group: $\mathrm{Trans}(\mathbf{R}^2)$

The group of rotations and reflections of $\mathbf{R}^2$ play the role as the set of all symmetries of $\mathbf{R}^2$ with its usual Euclidean geometry. This is one of the primary reasons mathematicians (or even physicists or computer scientists). In fact, these symmetries should really be called rigid symmetries since both reflections across lines through the origin or rotations about the origin, in both cases, take the origin to the origin; every linear function as defined here takes the origin to the origin. There is one addition symmetry that of $\mathbf{R}^2$ given by translation. Given a vector $v \in \mathbf{R}^2$, we define **translation by v** to be the function $T_v\colon \mathbf{R}^2 \to \mathbf{R}^2$ given by $T_v(x) = x+v$. Each translation T_v is completely determined by $T_v(0)$ where 0 is the zero vector. Under a translations, if we think of a vector x as a vector based as zero, one can view $T_v(x)$ as the vector x in the plane $\mathbf{R}^2$ but instead of starting at the origin, it starts at the point in $\mathbf{R}^2$ where the head of v ends. You image taking your regular directions of north, south, east, and west, and then translating them to point where v ends in $\mathbf{R}^2$. You would just first slide things left or right, and then slide them up or down. Just depends on what quadrant this point is in. It is worth noting that every translation $T_v\colon \mathbf{R}^2 \to \mathbf{R}^2$ is a bijective function. To see this, recall it is enough to produce an inverse function $f\colon \mathbf{R}^2 \to \mathbf{R}^2$ such that $(f \circ T_v)(x) = (T_v \circ f)(x) = x$ for all $x \in \mathbf{R}^2$. The inverse of T_v is easy to find. Just translate but now by $-v$. In fact, given $v,w \in \mathbf{R}^2$, we see that

$$T_v \circ T_w = T_w \circ T_v = T_{v+w}.$$

Translations by the zero vector is clearly the identity function. This shows $(T_v \circ T_{-v})(x) = (T_{-v} \circ T_v)(x) = x$.

To end with a brief summary:

- $\mathrm{Trans}(\mathbf{R}^2)$ is the set of functions from $\mathbf{R}^2$ to $\mathbf{R}^2$ given by the $T_v(x) = x+v$.
- We know that $\mathrm{Trans}(\mathbf{R}^2)$ is a subgroup of $\mathrm{Bi}(\mathbf{R}^2)$.
- There is a bijection between $\mathrm{Trans}(\mathbf{R}^2)$ and $\mathbf{R}^2$ via $v \to T_v$.
- $T_v \circ T_w = T_{v+w}$.
- $T_v \circ T_{-v} = T_0$.
- $T_v \circ T_w = T_w \circ T_v$.

If define $\Psi\colon \mathbf{R}^2 \to \mathrm{Trans}(\mathbf{R}^2)$ defined by $\Psi(v) = T_v$, we see that

$$\Psi(v+w) = T_{v+w} = T_v \circ T_w = \Psi(v) \circ \Psi(w).$$

In particular, Ψ is a homomorphism of groups where $\mathbf{R}^2$ is viewed a group under vector addition. Since Ψ is a bijection, this shows that $(\mathbf{R}^2,+) \cong \mathrm{Trans}(\mathbf{R}^2)$.

5.3.5 The affine group: $\mathrm{Aff}(\mathbf{R}^2)$

Using translations, we can define a group of functions $\mathrm{Aff}(\mathbf{R}^2)$ comprised of all invertible linear functions and all translations and call this the **affine group of $\mathbf{R}^2$**. We will think of these functions as a pair (v,A) where $A \in \mathrm{GL}(2,\mathbf{R})$ and $v \in \mathbf{R}^2$. The pair (v,A) will be thought of as a function $\mathbf{R}^2 \to \mathbf{R}^2$ given

$$(v,A)(x) = Ax+v.$$

Note that

$$(v,\mathrm{I}_2) = T_v, \quad (0,A) = A, \quad (v,A) = T_v \circ A.$$

We define $\mathrm{Aff}(\mathbf{R}^2)$ to be the set of all of these functions (v,A). Note that $\mathrm{Aff}(\mathbf{R}^2) \subset \mathrm{Fun}(\mathbf{R}^2)$. $\mathrm{Fun}(\mathbf{R}^2)$ is a monoid under composition of functions with identity being the identity function $\mathrm{Id}_{\mathbf{R}^2}$. Taking $A = \mathrm{I}_2$ and $v = 0$, we see that $(\mathrm{I}_2, 0)(x) = \mathrm{Id}_{\mathbf{R}^2}$. Given $(v,A),(w,B) \in \mathrm{Aff}(\mathbf{R}^2)$, we see that

$$(v,A)\circ(w,B)(x) = (v,A)(Bx+w) = A(Bx+w)+v = (AB)x+(Aw+v) = (Aw+v,AB).$$

This shows that $(v,A)\circ(w,B) \in \mathrm{Aff}(\mathbf{R}^2)$. We can now think of composition on $\mathrm{Fun}(\mathbf{R}^2)$ as a binary operator $\star$ on $\mathrm{Aff}(\mathbf{R}^2)$ defined by the explicit formula

$$(v,A)\star(w,B) = (v+Aw,AB).$$

Note that

$$(0,\mathrm{I}_2)\star(v,A) = (0+\mathrm{I}_2v,\mathrm{I}_2A) = (v,A)$$

and

$$(v,A)\star(0,\mathrm{I}_2) = (v+A0,A\mathrm{I}_2) = (v,A).$$

Hence, $(0,\mathrm{I}_2)$ is an identity for the binary operator $\star$. It remains to find the inverse of (v,A) in $\mathrm{Aff}(\mathbf{R}^2)$ in order to prove that $\mathrm{Aff}(\mathbf{R}^2)$ is a group. We will find the inverse of (v,A) by computation. We will assume that (w,B) is the inverse of (v,A). Now, we know that

$$(v,A)\star(w,B) = (v+Aw,AB).$$

If (w,B) is the inverse, then we must have

$$(v,A)\star(w,B) = (0,\mathrm{I}_2).$$

Combining these two facts, we see that

$$(v+Aw,AB) = (0,\mathrm{I}_2).$$

In particular, we have

$$v+Aw = 0, \quad AB = \mathrm{I}_2.$$

Solving for B in the second equation, we have $B = A^{-1}$. Solving for w in the first equation, we have

$$Aw = -v$$

and so

$$w = -A^{-1}v.$$

Hence, the inverse of (v,A) is $(-A^{-1}v,A^{-1})$. This shows that $\mathrm{Aff}(\mathbf{R}^2)$ is a group. Viewing $\mathrm{Aff}(\mathbf{R}^2) \subset \mathrm{Fun}(\mathbf{R}^2)$, the group operation is by composition of function and identity the identity function on $\mathbf{R}^2$. Viewing

$$\mathrm{Aff}(\mathbf{R}^2) = \mathbf{R}^2 \times \mathrm{GL}(2,\mathbf{R}) = \left\{(v,A) \;:\; v \in \mathbf{R}^2,\ A \in \mathrm{GL}(2,\mathbf{R})\right\},$$

the group operation is defined by $\star$ where

$$(v,A)\star(w,B) = (v+Aw,AB)$$

and identity $(0,\mathrm{I}_2)$. We can view $\mathrm{Aff}(\mathbf{R}^n)$ as all of the symmetries of $\mathbf{R}^2$ you get by allowing for rotations, reflections, scalings, and translates. Specifically, if we allow any finite combination of these operations, the set of all such functions or symmetries we can get is $\mathrm{Aff}(\mathbf{R}^2)$. It is the subgroup of $\mathrm{Bi}(\mathbf{R}^2)$ generated by those functions associated to the above operations. This group is called the **group of affine transformations of $\mathbf{R}^2$** or simply the **affine group for $\mathbf{R}^2$**. It is nice to have this geometric interpretation of $\mathrm{Aff}(\mathbf{R}^2)$. But if you had to take 15,000 elements $A_1,\dots,A_{15,000} \in \mathrm{Aff}(\mathbf{R}^2)$ and then take

$$(A_{15,000}\circ A_{14,999}\circ\cdots\circ A_3\circ A_2\circ A_1)(x)$$

you would likely use the binary operator $\star$ and program it. My point is that the formula for $\star$ gives you a way of quickly determining what combinations of the geometric operations do. It is hard to see the geometry in the formula but the formula really is taking all of those combinations of rations, reflections, scalings, and translations, and then packaging it into a fairly simple explicit binary operation a computer can easily handle. On the other hand, it hard to get intuition for what kind of functions comprise $\mathrm{Aff}(\mathbf{R}^2)$ through just the binary operator $\star$. You can visualize rotations, reflections, scalings, and translations. Both views are good.

5.3.6 Some explicit subgroups of $\mathrm{GL}(2,\mathbf{R})$: group of upper triangular matrices

Consider

$$G = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \ : \ a,b,d \in \mathbf{R},\ ad \neq 0 \right\}.$$

Note that $G \neq \emptyset$ since $a = d = 1$ and $b = 0$ gives $\mathrm{I}_2 \in G$. Given $A,B \in G$, we will check that $AB \in G$. Since $A,B \in G$, we know that

$$A = \begin{pmatrix} a_1 & b_1 \\ 0 & d_1 \end{pmatrix}, \quad B = \begin{pmatrix} a_2 & b_2 \\ 0 & d_2 \end{pmatrix}.$$

It follows that

$$AB = \begin{pmatrix} a_1 & b_1 \\ 0 & d_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ 0 & d_2 \end{pmatrix} = \begin{pmatrix} a_1a_2 & a_1b_2 + b_1d_2 \\ 0 & d_1d_2 \end{pmatrix} \in G.$$

Finally, given $A \in G$, we must check that $A^{-1} \in G$. As before, we can assume that

$$A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

and we want to find $B \in G$ such that $AB = \mathrm{I}_2$. We will assume that

$$B = \begin{pmatrix} x & y \\ 0 & z \end{pmatrix}.$$

Then we must have

$$AB = \begin{pmatrix} ax & ay + bz \\ 0 & dz \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{I}_2.$$

This gives four equations, one for each coefficient:

$$ax = 1, \quad ay + bz = 0, \quad 0 = 0, \quad dz = 1.$$

Since $a,d \neq 1$, we can solve for x,z in the first and last equations. We see then that $x = \frac{1}{a} = a^{-1}$ and $z = \frac{1}{d} = d^{-1}$. Inserting this into the second equation, we obtain

$$ay + \frac{b}{d} = 0.$$

Hence,

$$y = -\frac{b}{ad} = ba^{-1}d^{-1}.$$

Summary: We see that

$$G = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \ : \ a,b,d \in \mathbf{R},\ ad \neq 0 \right\}$$

is a group under matrix multiplication. For $A \in G$, we have

$$A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

where $a,d \in \mathbf{R}^\times$ and $b \in \mathbf{R}$. As an alternative, we can think of $G = \mathbf{R}^\times \times \mathbf{R}^\times \times \mathbf{R}$ where A is identified with (a,d,b). We can describe matrix multiplication for A,B in G where

$$A = \begin{pmatrix} a_1 & a_3 \\ 0 & a_2 \end{pmatrix} = (a_1,a_1,a_3), \quad B = \begin{pmatrix} b_1 & b_3 \\ 0 & b_2 \end{pmatrix} = (b_1,b_2,b_3)$$

as

$$AB = \begin{pmatrix} a_1b_1 & a_1b_3 + a_3b_2 \\ 0 & a_2b_2 \end{pmatrix} = (a_1b_1, a_2b_2, a_1b_3 + a_3b_2).$$

Hence, we can define a binary operator $\star$ on $G = \mathbf{R}^\times \times \mathbf{R}^\times \times \mathbf{R}$ defined by

$$(a_1, a_2, a_3) \star (b_1, b_2, b_3) = (a_1 b_1, a_2 b_2, a_1 b_3 + a_3 b_2).$$

For example if $A = (1,2,2)$ and $B = (3,5,-1)$, then

$$A \star B = (1,2,2) \star (3,5,-1) = (3,10,-1+10) = (3,10,9).$$

In matrix format, we see that

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & -1 \\ 0 & 5 \end{pmatrix}, \quad A \star B = \begin{pmatrix} 3 & 9 \\ 0 & 10 \end{pmatrix}.$$

Calculating AB, we see that

$$\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 3 & 9 \\ 0 & 10 \end{pmatrix}.$$

It is sometimes easier to write the coordinates for G as (x,y,z) and then write $\star$ as

$$(x,y,z) \star (x',y',z') = (xx', yy', xz' + zy').$$

The identity for $\star$ is $1_G = (1,1,0)$ and the inverse of (x,y,z) under $\star$ is given by $(x^{-1}, y^{-1}, zx^{-1}y^{-1})$. For instance, in this notation, setting $g = (x,y,z)$, we have

$$g^2 = (x,y,z) \star (x,y,z) = (x^2, y^2, xz + yz).$$

Next, we have

$$\begin{aligned} g^3 = (x,y,z) \star (x,y,z) \star (x,y,z) &= (x^2, y^2, z(x+y)) \star (x,y,z) \\ &= (x^3, y^3, x^2 z + yz(x+y)) \\ &= (x^3, y^3, x^2 z + xyz + y^2 z). \end{aligned}$$

Next, we have

$$\begin{aligned} g^4 = (x,y,z) \star (x,y,z) \star (x,y,z) \star (x,y,z) &= (x^3, y^3, x^2 z + y^2 z + xyz) \star (x,y,z) \\ &= (x^4, y^4, x^3 z + y(x^2 z + y^2 z + xyz)) \\ &= (x^4, y^4, x^3 z + x^2 yz + xy^2 z + y^3 z). \end{aligned}$$

One can deduce a pattern. Specifically,

$$g^n = (x^n, y^n, F_n(x,y,z))$$

where

$$F_n(x,y,z) = x^{n-1} z + x^{n-2} yz + x^{n-3} y^2 z + \cdots + x^2 y^{n-3} z + xy^{n-2} z + y^{n-1} z.$$

We can also write $F_n(x,y,z)$ in summation notation

$$F_n(x,y,z) = \sum_{i=0}^{n-1} x^{n-1-i} y^i z.$$

In particular, we see that

$$g^7 = (x^7, y^7, x^6 z + x^5 yz + x^4 y^2 z + x^3 y^3 z + x^2 y^4 z + xy^5 z + y^6 z).$$

Here are some additional, more explicit examples.

$$\begin{aligned} g &= (a,0,1), & g^n &= (a^n, a^{n-1}) \\ g &= (a,a^{-1},0), & g^n &= (a^n, a^{-n}, 0) \\ g &= (1,1,a), & g^n &= (1,1,an). \end{aligned}$$

5.3.7 Some subgroups of upper triangular matrices

Inside the group of upper triangular matrices

$$\left\{ \begin{pmatrix} x & z \\ 0 & y \end{pmatrix} \ : \ x,y,z \in \mathbf{R},\ xy \neq 0 \right\},$$

we have a a few subgroups, which we write in both matrix and coordinate form:

$$\begin{aligned}
H_1 &= \left\{ \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} \ : \ z \in \mathbf{R} \right\} = \{(1,1,z) \ : \ z \in \mathbf{R}\} \\
H_2 &= \left\{ \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \ : \ x,y \in \mathbf{R}^\times \right\} = \left\{(x,y,0) \ : \ x,y \in \mathbf{R}^\times\right\} \\
H_3 &= \left\{ \begin{pmatrix} x & z \\ 0 & 1 \end{pmatrix} \ : \ x \in \mathbf{R}^\times,\ z \in \mathbf{R} \right\} = \left\{(x,1,z) \ : \ x \in \mathbf{R}^\times,\ z \in \mathbf{R}\right\} \\
H_4 &= \left\{ \begin{pmatrix} x & 0 \\ 0 & x^{-1} \end{pmatrix} \ : \ x \in \mathbf{R}^\times \right\} = \left\{(x,x^{-1},0) \ : \ x \in \mathbf{R}^\times\right\} \\
H_5 &= \left\{ \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} \ : \ x \in \mathbf{R}^\times \right\} = \left\{(x,1,0) \ : \ x \in \mathbf{R}^\times\right\}.
\end{aligned}$$

We also write the binary operator $\star$ on each of these subgroups:

$$\begin{aligned}
H_1 : &\quad (1,1,z) \star (1,1,z') = (1,1,z+z') \\
H_2 : &\quad (x,y,0) \star (x',y',0) = (xx',yy',0) \\
H_3 : &\quad (x,1,z) \star (x',1,z') = (xx',1,xz'+z) \\
H_4 : &\quad (x,x^{-1},0) \star (x',(x')^{-1},0) = (xx',(xx')^{-1},0) \\
H_5 : &\quad (x,1,0) \star (x',1,0) = (xx',1,0).
\end{aligned}$$

The reader can see that $H_1 \cong (\mathbf{R},+)$, $H_2 \cong \mathbf{R}^\times \times \mathbf{R}^\times$, and $H_4, H_5 \cong \mathbf{R}^\times$. The group H_3 is more difficult to describe and is an example of a semidirect product.

Remark: The group $(\mathbf{R},+)$ will be denoted by $\mathbf{R}_a$ (also by $\mathbf{G}_{\mathrm{a}}(\mathbf{R})$) and call this the **additive group associated to R**. The group $(\mathbf{R}^\times,\cdot)$ will be denoted by $\mathbf{R}_m$ (also by $\mathbf{G}_{\mathrm{m}}(\mathbf{R})$) and is called the **multiplicative group associated to R**. These are the groups associated to $\mathbf{R}$ under addition $\mathbf{R}_a$ and multiplication $\mathbf{R}_m$. To illustrate the claimed relationships, let us consider $H_1 \cong \mathbf{R}_a = \mathbf{R}$. We see that

$$H_1 = \{(1,1,z) \ : \ z \in \mathbf{R}\}.$$

To prove that $H_1 \cong \mathbf{R}_a$, we must construct an isomorphism $\psi \colon H_1 \to \mathbf{R}_a = \mathbf{R}$. Explicitly, we need a bijective function $\psi \colon H_1 \to \mathbf{R}_a = \mathbf{R}$ such that

$$\psi((1,1,z) \star (1,1,z')) = \psi((1,1,z)) + \psi((1,1,z')).$$

Given $(1,1,z) \in H_1$, we define $\psi((1,1,z)) = z$. This is a bijective function $\psi \colon H_1 \to \mathbf{R}_a$. We must check that

$$\psi((1,1,z) \star (1,1,z')) = \psi((1,1,z)) + \psi((1,1,z')).$$

For that, we see that

$$(1,1,z) \star (1,1,z') = (1,1,z+z').$$

Therefore, by definition of ψ, we have

$$\psi((1,1,z) \star (1,1,z')) = \psi(1,1,z+z') = z+z'.$$

By definition of ψ, we have

$$\psi((1,1,z))+\psi((1,1,z'))=z+z'.$$

Hence, ψ is a bijective group homomorphism and so an isomorphism (by definition of isomorphism).

We will also workout the case $H_2 \cong \mathbf{R}_m \times \mathbf{R}_m$. We have

$$H_2=\left\{(x,y,0)\ :\ x,y\in\mathbf{R}^\times\right\}.$$

We define $\psi\colon H_2 \to \mathbf{R}_m \times \mathbf{R}_m = \mathbf{R}^\times \times \mathbf{R}^\times$ by

$$\psi((x,y,0))=(x,y).$$

The group operation on $\mathbf{R}_m \times \mathbf{R}_m$ is multiplication in each factor. That is $(x,y),(x',y')\in\mathbf{R}^\times\times\mathbf{R}^\times$, we have

$$(x,y)(x',y')=(xx',yy').$$

Therefore,

$$\psi((x,y,0))\psi((x',y',0))=(xx',yy').$$

We also know that

$$(x,y,0)\star(x',y',0)=(xx',yy').$$

Hence, by definition of ψ, we have

$$\psi((x,y,0)\star(x',y',0))=(xx',yy')=\psi((x,y,0))\psi((x',y',0)).$$

Therefore, ψ is a bijective group homomorphism and so H_2 and $\mathbf{R}_m\times\mathbf{R}_m$ are isomorphic.

Remark: We also have the subgroup

$$V_4=\left\{\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}-1&0\\0&1\end{pmatrix},\begin{pmatrix}1&0\\0&-1\end{pmatrix},\begin{pmatrix}-1&0\\0&-1\end{pmatrix}\right\}.$$

To further investigate V_4, let

$$\begin{aligned}1&=\begin{pmatrix}1&0\\0&1\end{pmatrix},\\ a&=\begin{pmatrix}-1&0\\0&1\end{pmatrix},\\ b&=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.\end{aligned}$$

Then $V_4=\{1,a,b,ab\}$ and

$$a^2=1,\quad b^2=1,\quad ab=ba,\quad (ab)^2=1.$$

So V_4 is a group under matrix multiplication and $|V_4|=4$. We can think of $V_4=\{1,a,b,c\}$ where multiplication $\cdot$ is defined by (this is like a multiplication table)

$$\begin{array}{llll}1\cdot1=1, & 1\cdot a=a, & 1\cdot b=b, & 1\cdot c=c,\\ a\cdot1=a, & a\cdot a=1, & a\cdot b=c, & a\cdot c=b,\\ b\cdot1=b, & b\cdot a=c, & b\cdot b=1, & b\cdot c=a,\\ c\cdot1=c, & c\cdot a=b, & c\cdot b=a, & c\cdot c=1.\end{array}$$

We see that $a^{-1}=a$, $b^{-1}=b$, and $c^{-1}=c$. The group V_4 is sometimes called the **Klein-4 group**.

5.3.8 Higher dimensions: Heisenberg group

We define

$$\mathrm{Heis}_3(\mathbf{R}) = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \ : \ x,y,z \in \mathbf{R} \right\}.$$

The group $\mathrm{Heis}_3(\mathbf{R})$ is called the **3–dimensional Heisenberg group**. Given $A,B \in \mathrm{Heis}_3(\mathbf{R})$, one can check that $AB \in \mathrm{Heis}_3(\mathbf{R})$. Specifically, we have

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & x' & z' \\ 0 & 1 & y' \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x+x' & z+z'+xy' \\ 0 & 1 & y+y' \\ 0 & 0 & 1 \end{pmatrix}.$$

Viewing

$$A = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

as $A = (x,y,z)$, we see that if $A' = (x',y'z')$, then the above matrix multiplication can be viewed as the binary operator $\star$ given by

$$A \star A' = (x,y,z) \star (x',y',z') = (x+x',y+y',z+z'+xy').$$

Let $X(x) = (x,0,0)$, $Y(y) = (0,y,0)$, and $Z(z) = (0,0,z)$. We will first compute the inverses of $X(x)$, $Y(y)$, and $Z(z)$.

We see that

$$X(x) \star Y(y) = (x,y,xy), \quad Y(x) \star X(x) = (x,y,0).$$

Hence, $[X(x),Y(x)] \neq 1$. To compute $[X(x),Y(y)]$, we compute X^{-1},Y^{-1}. We see that

$$X(x) \star X(-x) = (0,0,0)$$

and

$$Y(y) \star Y(-y) = (0,0,0).$$

Using this, we see that

$$\begin{aligned}(X(x))^{-1}(Y(y))^{-1}X(x)Y(y) &= (-x,0,0)(0,-y,0)(x,0,0)(0,y,0) = (-x,-y,xy)(x,y,xy) \\ &= (0,0,2xy-xy) = (0,0,xy) = Z(xy).\end{aligned}$$

Notice also that

$$Z(z) \star X(x) = X(x) \star Z(z), \quad Z(z) \star Y(y) = Y(y) \star Z(z).$$

Given $\gamma \in \mathrm{Heis}_3(\mathbf{R})$ with $\gamma = (x,y,z)$, we want to find $x',y',z' \in \mathbf{R}$ such that

$$\gamma = X(x')Y(y')Z(z').$$

For that, we see that

$$X(x')Y(y')Z(z') = (x',y',x'y')(0,0,z') = (x',y',z'+x'y').$$

Hence, $x' = x$, $y' = y$, and $z' = z - xy$. That is, $\gamma = X(x)Y(y)Z(z-xy)$.

Summary: We can think of $\mathrm{Heis}_3(\mathbf{R})$ as $\mathbf{R}^3$ via the coordinates $X(x),Y(y),Z(z)$. The group operation is given by

$$(X(x),Y(y),Z(z)) \star (X(x'),Y(y'),Z(z')) = (X(x+x'),Y(y+y'),Z(z+z'+xy')).$$

In particular, in the first two coordinates, the group operation is like $\mathbf{R}^2$ with vector addition. The third coordinate is not like addition though since it also has the xy' term. One can ask what functions $F\colon \mathbf{R}^6 = \mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$ work to give a group operation. Specifically, if we define

$$(x,y,z) \star_F (x',y',z') = (x+x', y+y', z+z'+F(x,y,z,x',y',z')),$$

we obtain a binary operation $\star_F$ on $\mathbf{R}^3$ that depends on the choice of the function F. For what choices of $F$is $\star_F$ a semigroup, monoid, or group operation? You could star with the case when F depends only on x, y' like in the Heisenberg group.

5.3.9 Higher dimensions: Orthogonal and special orthogonal groups

We define $\mathrm{O}(m) < \mathrm{GL}(m,\mathbf{R})$ by

$$\mathrm{O}(m) = \{A \in \mathrm{M}(n,\mathbf{R}) \ : \ Ax \cdot Ay = x \cdot y \text{ for all } x,y \in \mathbf{R}^m\}.$$

Proposition 5.10. *Let $A \in \mathrm{M}(m,\mathbf{R})$. Then the following are equivalent:*

(i) $A \in \mathrm{O}(m)$.

(ii) $AA^T = \mathrm{I}_m$.

(iii) The column vectors $a_1,\dots,a_m$ of A are an orthonormal set.

(iv) The row vectors $a^1,\dots,a^m$ of A are an orthonormal set.

Since $\det(A) = \det(A^T)$ and $\det(A)\det(A^T) = 1$, we see that if $A \in \mathrm{O}(m)$, then $(\det(A))^2 = 1$. Hence, $\det(A) = \pm 1$. We define

$$\mathrm{SO}(m) = \mathrm{O}(m) \cap \mathrm{SL}(m,\mathbf{R}).$$

W call $\mathrm{O}(m)$ the **orthogonal group** and $\mathrm{SO}(m)$ the **special orthogonal group**. Here are some extra references for the orthogonal and special orthogonal groups:

(1) $\mathrm{SO}(2)$

(2) $\mathrm{SO}(3)$

(3) $\mathrm{SO}(4)$

(4) $\mathrm{O}(2)$

(5) $\mathrm{O}(4)$ and the hydrogen atom

5.4 $\mathrm{Sym}(m)$: The symmetric group on m objects

Taking $X = \{1,\dots,m\}$, the group $\mathrm{Bi}(X)$ is typically called the **symmetric group** or **the group of permutations of m objects** and is denoted by $\mathrm{Sym}(m)$. These groups form an important class of finite groups and have been extensively studied. They appear throughout mathematics from algebraic geometry, Galois theory, and combinatorics.

5.4.1 Special elements

5.4.2 Decompositions of elements

5.4.3 Conjugacy classes, normalizers, and centralizers

5.4.4 The alternating group: $\mathrm{Alt}(m)$

5.4.5 Other special subgroups

Dihedral groups:

Symmetric subgroups and products of symmetric groups:

Transitive subgroups:

Cayley's Theorem:

5.4.6 One realization of the symmetric group by matrices

Taking $X = \{1,\dots,m\}$ and $\sigma \in \mathrm{Bi}(X)$, we define $L_\sigma \colon \mathbf{R}^m \to \mathbf{R}^m$ by

$$L_\sigma(v) = \sum_{i=1}^m \alpha_i e_{\sigma(i)}$$

where

$$v = \sum_{i=1}^m \alpha_i e_i.$$

From this, we obtain a homomorphism $\psi \colon \mathrm{Bi}(X) \to \mathrm{GL}(m,\mathbf{R})$ defined by $\psi(\sigma) = L_\sigma$. In fact, one can check that $L_\sigma \in \mathrm{O}(m)$.

5.5 Free groups

Given a set X, the **free group on the set X** is the unique group (up to isomorphism) $F(X)$ with $X \subset F(X)$ and satisfying the following universal mapping property: if $f \colon X \to G$ is any function where G is a group, then there exists a unique homomorphism $\varphi_f \colon F(X) \to G$ such that the restriction of φ_f to X is f. The existence of such a group can be done constructively in the same spirit as the construction of a vector space or free commutative group associated to a set. We briefly outline the approach.

We start by forming a set $\mathscr{A} = \{x\}_{x\in X} \cup \left\{x^{-1}\right\}_{x\in X}$ that we will call an **alphabet**. A **word** in the alphabet $\mathscr{A}$ is a finite ordered subset $w \subset \mathscr{A}$. By definition, $w = \left\{x_1^{\varepsilon_1},\dots,x_m^{\varepsilon_m}\right\}$ where $x_i \in X$ and $\varepsilon_i = \pm 1$. We can combine words $w_1 = \left\{x_1^{\varepsilon_1},\dots,x_m^{\varepsilon_m}\right\}$ and $w_2 = \left\{y_1^{\varepsilon_1'},\dots,y_n^{\varepsilon_n'}\right\}$ to form a new word $w_1w_2 \subset \mathscr{A}$ given by

$$w_1w_2 = \left\{x_1^{\varepsilon_1},\dots,x_m^{\varepsilon_m},y_1^{\varepsilon_1'},\dots,y_n^{\varepsilon_n'}\right\}.$$

We say that a word $w = \left\{x_i^{\varepsilon_i}\right\}_{i=1}^m$ is **reducible** if there exists an index $i \in \{1,\dots,n-1\}$ such that $x_i = x_{i+1}$ and $\varepsilon_i = -\varepsilon_{i+1}$, and we say w is **irreducible** otherwise. Given a reducible word $w \subset \mathscr{A}$ with $i \in \{1,\dots,n-1\}$ such that $x_i = x_{i+1}$ and $\varepsilon_i = -\varepsilon_{i+1}$, we can perform a reduction operation on w. Formally, we define a new word

$$\mathrm{R}_i(w) = \left\{x_1^{\varepsilon_1},\dots,x_{i-1}^{\varepsilon_{i-1}},x_{i+2}^{\varepsilon_{i+2}},\dots,x_n^{\varepsilon_n}\right\}.$$

After applying finitely many reduction operations $R_{i_1}, \ldots, R_{i_m}$ on the reducible word w, we obtain a unique irreducible word w_{irr}. We define $\mathscr{W}_X$ to be the set of irreducible words in the alphabet $\mathscr{A}$. We have a binary operation $*\colon \mathscr{W}_X \times \mathscr{W}_X \to \mathscr{W}_X$ given by $w_1 * w_2 = (w_1 w_2)_{irr}$. We also have an involution $\iota\colon \mathscr{W}_X \to \mathscr{W}_X$ given by

$$\iota(w) = \iota\left(\left\{x_i^{\varepsilon_i}\right\}_{i=1}^n\right) = \left\{x_{n-i+1}^{-\varepsilon_{n-i+1}}\right\}_{i=1}^n.$$

Setting $1_{\mathscr{W}_X} = \emptyset$, we see that $(\mathscr{W}_X, *, \iota, 1_{\mathscr{W}_X})$ is a group with multiplication operation $*$, inverse operation ι, and identity $1_{\mathscr{W}_X}$. Given any group G and any function $f\colon X \to G$, we define $\varphi_f\colon \mathscr{W}_X \to G$ by

$$\varphi_f(w) = \prod_{i=1}^n f(x_i)^{\varepsilon_i}, \quad w = \left\{x_i^{\varepsilon_i}\right\}_{i=1}^n.$$

It is straightforward to check that φ_f is a homomorphism with $f(x) = \varphi_f(x)$ for all $x \in X$.

6 Group Actions

6.1 Basics

Definition 67 (Group Action)**.** Given a group G and a set X, a **(left) group action** of G on X is a function $\varphi\colon G \times X \to X$ that satisfies the following two properties:

(a) For each $x \in X$, we have $\varphi(e,x) = x$.

(b) For each $g,h \in G$ and $x \in X$, we have $\varphi(gh,x) = \varphi(g,\varphi(h,x))$.

The following exercise is important in revealing the nature of what a group action is.

Exercise 1. *Let $\varphi\colon G \times X \to X$ be a group action on X and let $\mathrm{Bi}(X)$ denote the set of bijection function $\lambda\colon X \to X$.*

(i) Prove that $\mathrm{Aut}_{\mathrm{set}}(X)$ is a group where the identity element is given by the function $\mathrm{Id}_X : X \to X$ defined by $\mathrm{Id}_X(x) = x$ and the binary operation on $\mathrm{Aut}_{\mathrm{set}}(X)$ is composition of functions.

(ii) For each $g \in G$, define the function $\varphi_g : X \to X$ by $\varphi_g(x) = \varphi(g,x)$. Prove that $\varphi_g \in \mathrm{Aut}_{\mathrm{set}}(X)$.

(iii) Define the function $\Phi\colon G \to \mathrm{Aut}_{\mathrm{set}}(X)$ by $\Phi(g) = \varphi_g$. Prove that Φ is a homomorphism.

Exercise 2. *Prove that if $\Phi\colon G \to \mathrm{Sym}(X)$ is a homomorphism, then the function $\varphi\colon G \times X \to X$ given by $\varphi(g,x) = \Phi(g)(x)$ is a group action of G on X.*

In summary, Exercise 1 and Exercise 2 show that a group action of G on X is equivalent to a homomorphism $\Phi\colon G \to \mathrm{Sym}(X)$. Consequently, when we have a group action $\varphi\colon G \times X \to X$, we will simplify our notation and write $g \cdot x = \varphi(g,x)$. This notation is somewhat abusive and more precisely should be written as $\Phi(g)(x) = \varphi(g,x)$. However, it is extremely common to suppress the dependence on Φ so long as one is not considering several different actions of G on a fixed set X at once.

For the readers' clarity, we rewrite the definition of a group action in this simplified notation. A group action of G on X is a function $G \times X \to X$ denoted by $(g,x) \mapsto g \cdot x$ that satisfies the following properties:

(a) $e_G \cdot x = x$ for all $x \in X$ (i.e., e_G acts by the function Id_X).

(b) For each $g,h \in G$ and $x \in X$, we have $(gh) \cdot x = g \cdot (h \cdot x)$ (i.e., the group multiplication is the same as composition of functions).

We now discuss some basic examples of group actions. We start with one that we have already seen.

Example 1 (Symmetry Groups of Sets)**.** Given a set X, the group $\mathrm{Bi}(X)$ acts on X. We will prove that $(\lambda,x) \mapsto \lambda(x)$ is a group action. The identity element of $\mathrm{Bi}(X)$ is the identity function Id_X. We see that $\mathrm{Id}_X(x) = x$ and so property (a) for a group action holds. Likewise, given $\lambda_1, \lambda_2 \in \mathrm{Bi}(X)$ and $x \in X$, we see that

$$(\lambda_1\lambda_2) \cdot x \overset{\mathrm{def}}{=} (\lambda_1 \circ \lambda_2)(x) = \lambda_1 \cdot (\lambda_2 \cdot x).$$

Given a G–action on a set X, we next discuss the G–action on the space of complex valued function $f\colon X \to \mathbf{R}$. It is often the case that the set X is equipped with some additional structure, like a topology, and that the action of G on X is continuous/smooth/analytic with respect to this additional structure. In this case, the G–action on the space of functions will preserve the subspace of continuous/smooth/analytic functions. For simplicity, we will only consider the case when X is a set in the following example.

Example 2 (Function Spaces)**.** Let G be a group with an action on a set X. We can endow $\mathrm{Fun}(X)$ with a G–action via

$$(g\cdot f)(x) = f(g^{-1}\cdot x) \tag{10}$$

where $g \in G$, $f \in \mathrm{Fun}(X)$, and $x \in X$. Equivalently, for $g \in G$, we have the function $F_g \colon X \to X$ given by $F_g(x) = g\cdot x$, and define $g\cdot f \overset{\text{def}}{=} f\circ F_{g^{-1}}$. We will prove that this gives a G–action on $\mathrm{Fun}(X)$ so that the reader can see why the action is defined this way (i.e. why we take inverses). To see that the identity element of G acts as the identity, we have

$$(e_G\cdot f)(x) \overset{\text{def}}{=} f(e_G^{-1}\cdot x) = f(e_G\cdot x) = f(x).$$

Next, we check the compatibility condition, and must show that

$$((gh)\cdot f)(x) = (g\cdot(h\cdot f))(x).$$

To that end, we have

$$\begin{aligned}((gh)\cdot f)(x) &= f((gh)^{-1}\cdot x) = f((h^{-1}g^{-1})\cdot x)\\ &= f(h^{-1}\cdot(g^{-1}\cdot x) = (h\cdot f)(g^{-1}\cdot x) = (g\cdot(h\cdot f))(x).\end{aligned}$$

The action of G on $\mathrm{Fun}(X)$ is called the **contragradient action**.

Remark. For additional clarity, we discuss further why we must define the contragradient action as we did. If we replace (10) with

$$(g\cdot f)(x) = f(g\cdot x), \tag{11}$$

we see that

$$((gh)\cdot f)(x) = f((gh)\cdot x) = f(g\cdot(h\cdot x)) = (g\cdot f)(h\cdot x) = (h\cdot(g\cdot f))(x).$$

In general, $(h\cdot(g\cdot f))(x) \neq (g\cdot(h\cdot f))(x)$. Hence, (11) does not in general satisfy (b) in Definition 67.

We will relate the contragradient action of G on $\mathrm{Fun}(G)$ and the left action of G on itself in the next section. This relation will further illustrate why we define the contragradient action via (10).

Remark. The set $\mathrm{Fun}(X)$ is a vector space over $\mathbf{R}$. Scalar multiplication and vector addition are done point-wise via

$$(\alpha f)(x) = \alpha f(x), \quad (f_1+f_2)(x) = f_1(x)+f_2(x)$$

where $f_1, f_2 \in \mathrm{Fun}(X)$, $x \in X$, and $\alpha \in \mathbf{R}$. If $g \in G$, we see that

$$(g\cdot(\alpha f))(x) = (\alpha f)(g^{-1}\cdot x) = (\alpha(g\cdot f))(x)$$

and

$$(g\cdot(f_1+f_2))(x) = (f_1+f_2)(g^{-1}\cdot x) = f_1(g^{-1}\cdot x)+f_2(g^{-1}\cdot x) = ((g\cdot f_1)+(g\cdot f_2))(x).$$

In particular, the function $T_g \colon \mathrm{Fun}(X) \to \mathrm{Fun}(X)$ define by $T_g(f) = g\cdot f = f\circ F_{g^{-1}}$ is a linear function. Let $\mathrm{Aut}_{\mathrm{vec}}(\mathrm{Fun}(X))$ be the set of bijective linear functions $T\colon \mathrm{Fun}(X) \to \mathrm{Fun}(X)$. We can endow $\mathrm{Aut}_{\mathrm{vec}}(\mathrm{Fun}(X))$ with a group structure where the identity element is the identity function and the group operation is composition of functions. The contragradient action of G on $\mathrm{Fun}(X)$ induces a group homomorphism $\Phi_{\mathrm{contra}}\colon G \to \mathrm{Aut}_{\mathrm{vec}}(\mathrm{Fun}(X))$. Specifically, $\Phi_{\mathrm{contra}}(g) = T_g$.

Definition 68 (Transitive Action)**.** Let G be a group with an action on a set X. We say that G acts **transitively** on X if for each pair $x_1, x_2 \in X$, there exists $g \in G$ such that $g\cdot x_1 = x_2$.

One often views X as a space/universe and in this view, a transitive G–action on X is a G–action in which one can go from any point in X to any other point in X via an application of an element of G. The group $\mathrm{Bi}(X)$ acts transitively on the set X. In fact, this action is highly transitive in the following sense. Given any subsets $S_1, S_2 \subset X$ with $|S_1| = |S_2|$, there exists $\sigma \in \mathrm{Bi}(X)$ such that $\sigma(S_1) = S_2$. One the other hand, if we take the subgroup of $\mathrm{Bi}(X)$ of all elements that fix $x_0 \in X$ (i.e. $\sigma(x_0) = x_0$), this subgroup of $\mathrm{Bi}(X)$ does not act transitively on X; it does act transitively on $X - \{x_0\}$.

Definition 69 (Faithful Action)**.** Let G be a group with an action on a set X. We say that G acts **faithfully** on X if for each non-trivial $g \in G$, there exists $x \in X$ such that $g \cdot x \neq x$.

The following lemma shows that faithful actions arise precisely from injective homomorphisms $\Phi \colon G \to \mathrm{Bi}(X)$.

Lemma 6.1. *Let G be a group with an action on a set X. Then the following are equivalent:*

(a) G acts faithfully on X.

(b) The associated homomorphism $\Phi \colon G \to \mathrm{Bi}(X)$ is injective.

If G is a group with an action on a set X, we define two basic subsets of G and X, respectively. For any subset $S \subseteq X$, we define

$$\mathrm{Stab}_G(S) = \{g \in G \;:\; g \cdot s \in S \text{ for all } s \in S\}$$

and

$$\mathscr{O}_{G,S} = \{g \cdot s \;:\; g \in G,\; s \in S\}\,.$$

We call $\mathrm{Stab}_G(S)$ the **stabilizer** of S and $\mathscr{O}_{G,S}$ the **orbit** of S (see also **here**).

Exercise 3. *Prove that if G is a group with an action on X and $S \subseteq X$, then $\mathrm{Stab}_G(S) \leq G$.*

The concept of a free action is a strengthening of a faithful action.

Definition 70 (Free Action)**.** Let G be a group with an action on a set X. We say that the action of G on X is **free** if for each $x \in X$, $\mathrm{Stab}_G(x) = \{e\}$.

Exercise 4. *Let G be a group with an action on a set X and associated homomorphism $\Phi \colon G \to \mathrm{Sym}(X)$.*

(i) Prove that $g \in \ker \Phi$ if and only if $g \in \mathrm{Stab}_G(x)$ for all $x \in X$.

(ii) Deduce that if G acts freely on X, then G acts faithfully on X.

(iii) Prove that $\mathrm{Sym}(X)$ acts freely on X if and only if $|X| \leq 2$. In particular, a faithful action need to be free.

Exercise 5. *Let G be a group with an action on X with associated homomorphism $\Phi \colon G \to \mathrm{Sym}(X)$. Prove that*

$$\bigcap_{x \in X} \mathrm{Stab}_G(x) = \ker \Phi.$$

6.2 The Orbit-Stabilizer Theorem

We now state a basic result for group actions that is often referred to as the **Orbit-Stabilizer Theorem**.

Theorem 6.2 (Orbit-Stabilizer Theorem)**.** *Let G be a group, X a set with a G–action, and $x \in X$.*

(a) For each $x_1, x_2 \in \mathscr{O}_x$, there exists $g_{2,1} \in G$ such that $\mathrm{Stab}_G(x_2) = g_{2,1}^{-1} \mathrm{Stab}_G(x_1) g_{2,1}$.

(b) There exists a bijective function $\lambda \colon G/\mathrm{Stab}_G(x) \to \mathscr{O}_x$.

(c) If G acts transitively on X, then there exists a bijective function $\lambda \colon G/\mathrm{Stab}_G(x) \to X$.

Proof. For (a), since $x_1, x_2 \in \mathscr{O}_x$, there exists $g_1, g_2 \in G$ such that $g_1 \cdot x = x_1$ and $g_2 \cdot x = x_2$. In particular, $g_1 g_2^{-1} \cdot x_2 = x_1$. Set $g_{2,1} = g_1 g_2^{-1}$. Given $g \in g_{2,1}^{-1} \operatorname{Stab}_G(x_1) g_{2,1}$. Then $g = g_{2,1}^{-1} g_0 g_{2,1}$ for some $g_0 \in \operatorname{Stab}_G(x_1)$. We have

$$\begin{aligned}(g_{2,1}^{-1} g_0 g_{2,1}) \cdot x_2 &= (g_{2,1}^{-1} g_0) \cdot (g_{2,1} x_2) = (g_{2,1}^{-1} g_0) \cdot x_1 \\ &= g_{2,1}^{-1} \cdot (g_0 x_1) = g_{2,1}^{-1} \cdot x_1 = x_2.\end{aligned}$$

Hence, $g \in \operatorname{Stab}_G(x_2)$. Given $g \in \operatorname{Stab}_G(x_2)$, it follows that

$$g = g_{2,1}^{-1} (g_{2,1} g g_{2,1}^{-1}) g_{2,1}.$$

We assert that $g_{2,1} g g_{2,1}^{-1} \in \operatorname{Stab}_G(x_1)$. To see this, we have

$$\begin{aligned}(g_{2,1} g g_{2,1}^{-1}) \cdot x_1 &= (g_{2,1} g) \cdot (g_{2,1}^{-1} x_1) = (g_{2,1} g) \cdot x_2 \\ &= g_{2,1} \cdot (g \cdot x_2) = g_{2,1} \cdot x_2 = x_1.\end{aligned}$$

Setting $g_0 = g_{2,1} g g_{2,1}^{-1}$, we see that $g = g_{2,1}^{-1} g_0 g_{2,1} \in g_{2,1}^{-1} \operatorname{Stab}_G(x_1) g_{2,1}$. We summarize pictorially the process of conjugating $\operatorname{Stab}_G(x_1)$ and $\operatorname{Stab}_G(x_2)$:

$$\begin{tikzcd}
x_1 \arrow[loop left, "\operatorname{Stab}_G(x_1)"] \arrow[r, bend left, "g_1^{-1}"] \arrow[rr, bend left=60, "g_{2,1}^{-1}=g_2g_1^{-1}"] & x \arrow[l, bend left, "g_1"] \arrow[r, bend left, "g_2"] & x_2 \arrow[loop right, "\operatorname{Stab}_G(x_2)"] \arrow[l, bend left, "g_2^{-1}"] \arrow[ll, bend left=60, "g_{2,1}=g_1g_2^{-1}"]
\end{tikzcd}$$

In the notation of the proof of (a), we also have the diagram with specific elements in place of the stabilizers:

$$\begin{tikzcd}
x_1 \arrow[loop left, "g_0"] \arrow[r, bend left, "g_1^{-1}"] \arrow[rr, bend left=60, "g_{2,1}^{-1}=g_2g_1^{-1}"] & x \arrow[l, bend left, "g_1"] \arrow[r, bend left, "g_2"] & x_2 \arrow[loop right, "g"] \arrow[l, bend left, "g_2^{-1}"] \arrow[ll, bend left=60, "g_{2,1}=g_1g_2^{-1}"]
\end{tikzcd}$$

For (b), we define $\lambda(g \operatorname{Stab}_G(x)) = g \cdot x$. To show that λ is well defined, we must show that if $g' \in g \operatorname{Stab}_G(x)$, then $g' \cdot x = g \cdot x$. By definition, $g' = g g_0$ where $g_0 \in \operatorname{Stab}_G(x)$. In particular, $g' \cdot x = (g g_0) \cdot x = g \cdot (g_0 \cdot x) = g \cdot x$, as needed. To prove that λ is bijective, we will prove that it is both injective and surjective. If $\lambda(g \operatorname{Stab}_G(x)) = \lambda(g' \operatorname{Stab}_G(x))$, we must show that $g \operatorname{Stab}_G(x) = g' \operatorname{Stab}_G(x)$. By definition, we have

$$g \cdot x = \lambda(g \operatorname{Stab}_G(x)) = \lambda(g' \operatorname{Stab}_G(x)) = g' \cdot x.$$

In particular, $g' g^{-1} \cdot x = x$ and so $g' g^{-1} \in \operatorname{Stab}_G(x)$. Hence $g \operatorname{Stab}_G(x) = g' \operatorname{Stab}_G(x)$. For surjectivity, given $x' \in \mathscr{O}_x$, we must find $g \operatorname{Stab}_G(x) \in G/\operatorname{Stab}_G(x)$ such that $\lambda(g \operatorname{Stab}_G(x)) = x'$. Since $x' \in \mathscr{O}_x$, there exists $g \in G$ such that $g \cdot x = x'$. By definition of λ, we see that $\lambda(g \operatorname{Stab}_G(x)) = x'$.

Part (c) follows immediately from (b) as $X = \mathscr{O}_x$ for any $x \in X$ when G acts transitively. □

Exercise 6. *Let G be a group which acts on X and $x \in X$ be fixed.*

(i) Prove that the function $\lambda_x \colon G \to X$ given by $\lambda_x(g) = g \cdot x$ is a bijective function if and only if G acts freely and transitively on X.

(ii) Assume that X is finite and $H \leq \operatorname{Sym}(X)$ acts freely and transitively on the set X. Prove that H is a cyclic group and $|H| = |X|$. Deduce that $\operatorname{Sym}(X)$ is not cyclic provided $|X| \geq 3$.

Exercise 7. *Let G be a group and X a set with a G–action. We define an equivalence relation on X as follows. Given $x, y \in X$, we say $x \sim_G y$ if and only if there exists $g \in G$ such that $g \cdot x = y$.*

(i) Prove that $\sim_G$ is an equivalence relation on X.

(ii) Prove that the set

$$[x]_G \overset{def}{=} \{y \in X \ : \ x \sim_G y\}$$

is equal to $\mathscr{O}_x$.

(iii) Using the axiom of choice, prove that there exists a subset $S \subset X$ such that for each $y \in X$, there exists a unique $x \in S$ such that $x \sim_G y$.

(iv) Prove that

$$X = \bigcup_{x \in S} \mathscr{O}_x.$$

(v) Deduce that there is a bijection between X and the set

$$\bigsqcup_{x \in S} G/\operatorname{Stab}_G(x)$$

where $\sqcup$ denotes the ***disjoint union****.*

Given sets X, Y, each with a G–action, we say that a function $f \colon X \to Y$ is ***G*–equivariant** if for each $x \in X$ and $g \in G$, we have $f(g \cdot x) = g \cdot f(x)$. After discussing some actions of G on itself, we will see that the function in the orbit stabilizer theorem $G/\operatorname{Stab}_G(x) \to \mathscr{O}_x$ is a G–equivariant map.