Notes on MAT1101 - Groups and Vector Spaces

Jake Xuereb

Last Updated - June 20, 2018

Preface

These notes will be based off of the lectures from MAT1101 given by Prof.Irene Scriha at the University of Malta within the Winter and Spring of 2018. These notes will be heavily involved with the material found in the book by Herstein, Topics in Algebra for Groups and the Cameron's Introduction to Algebra.

The **motivation** behind these notes is to recapitualise in a more formal manner what is discussed in the lectures and to compile information on the topics from multiple sources making for a more complete reference. These notes are an exercise of my personal thoughts and should in no way be considered official or affiliated with the University other than the fact that I attended these lectures. *Any comments or error-pointing related to these notes are to be directed at* **jqed.xuereb@gmail.com**

Contents

1 Introductory Considerations **3**

1.1 Set Relations . . . 3

1.1.1 The Cartesian Product . . . 3

1.1.2 Binary Relations . . . 3

1.1.3 Equivalence Relation . . . 4

1.1.4 Equivalence Classes and Partitions . . . 5

1.1.5 Congruence Modulo and Congruence Classes . . . 8

1.1.6 Posets . . . 8

1.2 Functions - taken and improved on from my notes on MAT1100 . . . 9

1.2.1 Surjectivity & Injectivity . . . 10

1.2.2 Bijections . . . 11

1.2.3 Compositions . . . 11

1.2.4 Inverse Functions . . . 11

1.2.5 Character Persistence in Composition . . . 12

1.2.6 A function of Symmetry ! . . . 12

2 Groups **13**

2.0.1 Preliminary Lemmas . . . 13

2.1 Notable Groups . . . 15

2.1.1 Formalisms . . . 15

2.1.2 Permutation Groups . . . 17

2.1.3 Visualising Groups through Symmetries . . . 18

2.2 Subgroups . . . 19

2.2.1 A Criterion for Subgroups . . . 19

2.2.2 Some Examples of Subgroups . . . 20

2.2.3 Cosets . . . 23

2.2.4 Lagrange's Theorem . . . 26

2.2.5 Corollaries of Lagrange's Theorem related to Elemental Order . . . 29

2.2.6 Corollaries of Lagrange's Theorem related to Number Theory . . . 31

2.3 Normal Subgroups . . . 38

2.3.1 Visualising Normal Subgroups using D_3 . . . 38

2.3.2 A Mathematical Definition of Normal Subgroups and Lemmas . . . 40

2.4 Homomorphisms and Quotient Groups . . . 45

2.4.1 Quotient Structures . . . 45

2.4.2 Homomorphisms 48
2.4.3 Noether's First Isomorphism Theorem 57
2.4.4 Cayley's Theorem 60

3 Vector Spaces **64**
3.1 Introduction and Formalisms 64
3.1.1 Examples 66
3.1.2 Properties of Vector Spaces 66
3.2 Linear Dependence, Span and Basis 68
3.2.1 Linear Dependence 68
3.2.2 Spanning Sets and Basis 70
3.3 Steinitiz Replacement 72
3.3.1 Further Basis and Dimension 74
3.4 The Dimension Theorem 76
3.5 Linear Transformations 76

Chapter 1

Introductory Considerations

1.1 Set Relations

1.1.1 The Cartesian Product

Firstly an **ordered pair** must be defined as a pair of objects were order is relevant $\implies (x, y)$ where x is always the first object and y is the second object.

$$(x, y) = (z, w) \iff x = z \land y = w$$

The **Cartesian Product** of two sets is the set of all ordered pairs where the first element always comes from the first set X and the second element of the ordered pairs always comes from the second set Y

$$X \times Y = \{(x, y) : x \in X, y \in Y\}$$

The need for a pair comes from the definition employed above making use of two sets. This notion may be generalised by using tuples as they are the mathematical structure which give importance to order. A practical example of a cartesian product is a deck of card where the two sets are the set of ranks containting 13 elements $\{2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A\}$ and the set of 4 suits $\{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$. The cartesian product of these two sets has 52 elements which are all ordered pairs.
Some examples for the cartesian product Rank $\times$ Suits include:

$$\{8, \spadesuit\}, \{4, \heartsuit\}, \{K, \diamondsuit\}$$

1.1.2 Binary Relations

The notion of a binary relation arises from as a subset of a **Cartesian Product** of two sets.

$$A \subseteq A \times B$$

Sticking to the earlier notion a subset of the deck of cards cartesian products would be the cartesian product Rank $\times$ Spades which would give

$$\{2,\spadesuit\},\{3,\spadesuit\},\{4,\spadesuit\},\{5,\spadesuit\},\{6,\spadesuit\},\{7,\spadesuit\},\{8,\spadesuit\},\{9,\spadesuit\},\{10,\spadesuit\},\{J,\spadesuit\},\{Q,\spadesuit\},\{K,\spadesuit\},\{A,\spadesuit\}$$

Now perhaps more pertinently if the cartesian product is a self product such that $(a,b) \in A \times A$ then we can start doing some fun stuff and talk about binary relations defined on A in terms of the elements of the set A.
So the way this is typically encountered would be in that a set is presented say $\mathbb{Z}$ and a relationship to be exhibited by the elements of the set is given as a conditional ex: if $a-b$ is an even integer then the **binary relation** $a \sim b$ is defined, where $\sim$ is the notation used for binary relations.

1.1.3 Equivalence Relation

This is a special collection of ordered pairs which are the subset of a cartesian self product that satisfy the properties

- Reflexitivity - miegħu nnifsu
- Symmetry - huma jimplikaw lilhom infushom bil-maqlub
- Transitivity - ovja eh

For $R \subseteq A \times A$ these properties would look like this in set notation in the order presented above

1. $(a,a) \in R \; \forall a \in A$
2. $(a,b) \in R \implies (b,a) \in R$
3. $(a,b) \in R \wedge (b,c) \in R \implies (a,c) \in R$

It is good to note here that it is the subset which possess the property or relation in question and thus in this view being a part of the subset R would show that said ordered pair possess the property which indeed makes R a subset of $A \times A$ and not $A \times A$ or some other subset possessing some other property.
If this subset, which represents a particular relation, satisfies these properties then this relation is an **equivalence relation**.

In terms of a binary relation $\sim$ on A

1. $a \sim a$
2. $a \sim b \implies b \sim a$
3. $a \sim b \wedge b \sim c \implies a \sim c$

The use of the binary relation allows one to describe the properties of a relation throughout a whole set making use of the elements of that set.

An *example*

For $a, b \in \mathbb{Z}, a \sim b$ if $a - b \in 2\mathbb{Z}$ (even). Show this is an equivalence relation.

$$a - a = 0$$ i) Reflexivity is shown as 0 is an even number

$$a \sim b \implies a - b \in 2\mathbb{Z}$$

Consider $b \sim a = b - a = -(a - b)$

$$\because a - b \in 2\mathbb{Z} \implies -(a - b) \in 2\mathbb{Z}$$ ii) Symmetry is shown

$$a - b \in 2\mathbb{Z} \wedge b - c \in 2\mathbb{Z}$$

Consider $a - c = (a - b) + (b - c) \in 2\mathbb{Z}$ iii) Transitivity is shown

Therefore all three axioms of the equivalence relation are upheld

1.1.4 Equivalence Classes and Partitions

Say $\sim$ on R is an *equivalence relation*. Think about how this relation involves always an ordered pair involving two numbers, a and b. Now given that this is an equivalence relation these numbers can take up specific forms with respect to the *relation* in question such that they satisfy the relation in a similar way.A different formulation of this idea would be, in what different ways can this relation be satisfied within this set R. These different ways or specific forms are called equivalence classes and are represented by the set

$$\text{cl}(a) = \{x \in R | a \sim x\}$$

So within the context of the example given above the numbers in the ordered pair have the relationship

$$a - b = 2m : m \in \mathbb{Z}$$

This is the form which the equivalence classes can take up. Reformulating this we get

$$a - 2m = b : m \in \mathbb{Z}$$

So a is either 0 given an **even** number or 1 to give an **odd** number since any other number is just a scaled version of these equivalence classes. Since a can only distinctly

take up the values of cl(0) and cl(1) then this equivalence relation has two distinct equivalence classes $cl0)$ and cl(1).

Further example : Consider $n : n \in \mathbb{Z} \wedge n > 1$. Define for $a, b \in \mathbb{Z}, a \sim b$ if $n|(a-b)$

i) For Reflivity consider $a - a = 0 \because w|0 : w \in \mathbb{Z}\square$

ii) For Symmetry consider $n|(a-b)$

$\implies b \sim a = b - a = -(a-b) \implies n| - (a-b)\square$

iii) For Transitivity consider $a - c = (a-b) + (b-c)$

But since the relation stands for $a \sim b \wedge a \sim c$

$\implies n|(a-b) + n|(b-c)$

$\implies n|(a-b+b-c) = n|(a-c)\square$

Therefore the relation $a \sim b$ is an equivalence relation on $\mathbb{Z}$.

Now the set at hand is that of the integers and we need to think about how many different ways this relation can be satisfied within the integers. The relation may be represented

$$a + kn : k \in \mathbb{Z}$$

Any a would create an equivalence class given that $a \neq n$

$$\implies \text{cl}(0), \text{cl}(1), \text{cl}(2), ..., \text{cl}(n-1)$$

Theorem 1.1.1. *The distinct equivalence classes of the equivalence relation on A allow for the decomposition of A as the union of mutually disjoint sets.*

Meaning that these equivalence classes can cover a whole set and are non-intersecting. These notions of subsets which are disjoint and whose union is the whole set; define the mathematical concept of a **partition** and this theorem is essentially saying that the equivalence classes of an equivalence relation on R partition R.

Proof. Given that two qualities must be satisfied the proof will follow directly in steps.

$$\begin{aligned}
&\text{i) We show } \bigcup_{i=1}^{r} \mathrm{cl}(a_i) = A \\
&\qquad \mathrm{cl}(a_i) \subseteq A && \text{by defn' of an equivalence class} \\
&\implies \bigcup_{i=1}^{r} \mathrm{cl}(a_i) \subseteq A \\
&\qquad \text{Reflexivity} : \forall a \in A, a \sim a && \because \sim \text{is an equivalence relation on A} \\
&\implies \forall a \in A, a \in \mathrm{cl}(a) \subseteq \bigcup_{i=1}^{r} \mathrm{cl}(a_i) \\
&\qquad \implies A \subseteq \bigcup_{i=1}^{r} \mathrm{cl}(a_i) \\
&\because \bigcup_{i=1}^{r} \mathrm{cl}(a_i) \subseteq A \wedge A \subseteq \bigcup_{i=1}^{r} \mathrm{cl}(a_i) \\
&\qquad \therefore \bigcup_{i=1}^{r} \mathrm{cl}(a_i) = A \ \square \\
&\text{ii) We show } \mathrm{cl}(a_i) \cap \mathrm{cl}(a_j) = \emptyset : i \neq j, 1 \geq i, j \geq r \\
&\qquad \text{Let } a \in \mathrm{cl}(a_i) \cap \mathrm{cl}(a_j) \\
&\qquad \iff a \in \mathrm{cl}(a_i) \wedge a \in \mathrm{cl}(a_j) && \text{by def'n of } \cap \\
&\qquad \implies a \sim a_i \wedge a \sim a_j && \text{by def'n of cl()} \\
&\qquad \implies a_i \sim a_j && \text{by transitivity} \\
&\qquad \text{Suppose } b \in \mathrm{cl}(a_i) \\
&\qquad \implies b \sim a_i && \text{by def'n of cl()} \\
&\qquad \implies b \sim a_j && \text{by } a_i \sim a_j \\
&\qquad \implies b \in \mathrm{cl}(a_j) && \text{by def'n of cl()} \\
&\qquad \therefore \mathrm{cl}(a_i) \subseteq \mathrm{cl}(a_j) && \text{by def'n of } \subseteq
\end{aligned}$$

The argument is symmetric and can be similarly shown for $\mathrm{cl}(a_j) \subseteq \mathrm{cl}(a_i)$

$$\therefore \mathrm{cl}(a_i) = \mathrm{cl}(a_j) \square$$

$\square$

1.1.5 Congruence Modulo and Congruence Classes

An excellent example of equivalence relations and equivalence classes is the congruence modulo.

The Congruence Modulo is a way of of representing numbers using divisors such that $\mathbf{a \equiv b} \mod \mathbf{n} : \mathbf{n}|(\mathbf{a - b})$, n is called the modulus of the relation.
For Example - $73 \equiv 4 \mod 23 : 23|(69)$

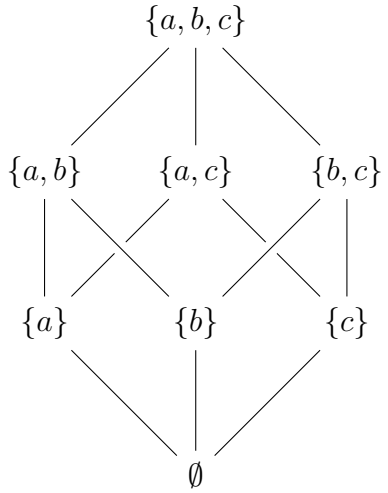
On page 6 of these notes it was shown that the congruence modulo is an equivalence relation on the set of integers. The equivalence classes of this relation are known as congruence classes and are denoted $[a]$.

1.1.6 Posets

A Partially Ordered Set or Poset is a binary relation on a set A with the following characteristics

- Reflexitivity - $a \sim a$
- Transitivity - $a \sim b \wedge b \sim c \implies a \sim c$
- Anti-Symmetry - $a \sim b \wedge b \sim a \implies a = b$

As shown below a Poset may be visualised or represented using a Hasse Diagram where in a set is in a sense decomposed by a relation where each line represents a relation. Within Hasse Diagrams objects on the same level are not comparable. Here under the relation being represented is $(S, \supseteq)$ where going through a whole path it is evident that $\emptyset$ is contained in all the sets above it and that $\{a\}$ is contained in $\{a, b\}$ and $\{a, b\}$ and finally $\{a, b, c\}$.
Other examples include divisor relations or $\geq$ relations.

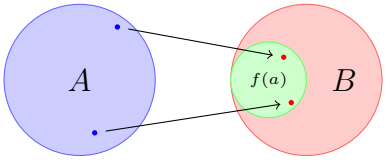


1.2 **Functions -** taken and improved on from my notes on MAT1100

A **function** is a set relation from a set A to a set B which to satisfy a rigorous definition must satisfy two conditions. Using the notation $f : A \to B$ we indicate that

$$i)\, \forall a \in A\ \exists\, b \in B : (a, b) \in f$$

meaning that each element in A is assigned an element from B of the form $f(a)$ and the region $f(a)$ is known as the image of a under f also referred to as the function. The domain of the function is A and the target or codomain is B.



$$f : A \to B$$

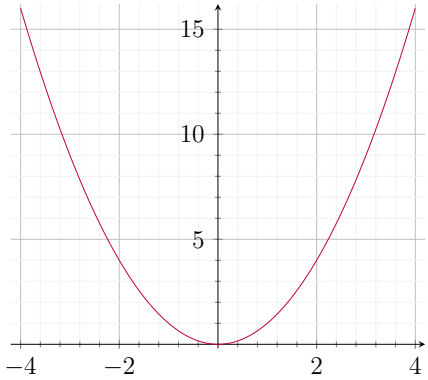
$$ii) \forall a \in A \wedge \forall b \in B,\ \text{if } (a, b) \in f \wedge (a, c) \in f \implies b = c$$

The second condition shown above is what is meant when one posits that a function is well-defined. For this condition to be satisfied to each element of A precisely one element of B is assigned. Thus each element of A is related to not more than one element of B.

Given that a function is a set relation between to sets it can be viewed as a subset of the cartesian product of the domain and the codomain. Within this same light one can think of it as enforcing a rule ex: $y = x^2$ enforcing the rule that two variables are related by one's square.

$$\implies f : A \to B \subseteq A \times B$$

The graph of a function $f : A \to B$ is $\{(x, f(x)) : x \in A\}$ this is the pictorial representation of this rule that we are all familiar with, for this particular example of $y = x^2$ we know it would like something like this



For a function $f : X \to Y$ the image of T under f, $f(T)$ such that $T \subseteq X$ is the set

$$f(T) = \{y \in Y : \exists x \in T : f(x) = y\}$$

The **inverse image** of a function is denoted by f^{-1} and is not to be confused with the inverse function. The inverse image applies the same notion of the image but applying it to a subset of the codomain of the function. Considering $C \subseteq Y$ the set will look something like this

$$f^{-1}(C) = \{x \in X : \exists y \in C : f(x) = y\}$$

1.2.1 Surjectivity & Injectivity

The section title includes two fancy names which describe in what way the domain is being mapped to the codomain and so what type of range the function creates.

A One-to-One function is one for which

$$\text{if } f(a_1) = f(a_2) \text{ then } a_1 = a_2$$

meaning that as the name implies there is a one to one assignment between the domain elements and the related image elements. These of functions are also known as injective.In terms of cardinality, for a function $f : A \to B$, if such a function is injective then this must imply that

$$|A| \leq |B|$$

and this makes intuitive sense as if every element of A corresponds to an element in the image which is a subset of B, then the totality of B is larger than A

An on-to function is one for which the range is equal to the codomain so for

$$f : A \to B \implies ran(f) = B$$
$$\forall y \in B \, \exists \, x \in A : f(x) = y$$

This implies that every element of the domain is related to an element of the codomain. These of functions are also known as surjective.In terms of cardinality, for a function $f : A \to B$, if such a function is surjective then this must imply that

$$|A| \geq |B|$$

and this makes intuitive sense because if each element of A is mapped to every element in B then it stands to reason that it is either a many to one or a one to one correspondence.

It is good to appreciate that an injective relation is a forward one going from the domain to the image and can be approached as such whilst a surjective relation is a backwards kind of thing where one starts out talk about the range and moves back to the domain.

1.2.2 Bijections

A bijective function is one which is both injective and surjective. Informally this may be referred to as a one to one correspondance between sets and can be expressed

$$\forall b \in B \ \exists \text{ precisely one } a \in A : f(x) = b$$

This infers that every element in the domain corresponds to single element in the range.

If one is able to form a bijection between two sets then it becomes clear that these two sets possess the same cardinality or size.

$$|A| = |B|$$

1.2.3 Compositions

Consider two functions f and g

$$f : A \to B \qquad\qquad g : C \to D$$

Their composition from A to D is the function $g \circ f : A \to D \equiv g(f(x))$

The composition of functions is associative but not commutative.

$$h \circ (g \circ f) = (h \circ g) \circ f \qquad\qquad f \circ g \neq g \circ f$$

Proof: **Associative Law :-** $h \circ (g \circ f) = (h \circ g) \circ f$

$$\begin{aligned}
\text{Restating } f : A \to B, g : B \to C \text{ and } h : C \to D &\implies h \circ (g \circ f) = (h \circ g) \circ f \\
\text{Consider } g \circ f : A \to C &\implies h \circ (g \circ f) : A \to D \\
\text{Consider } h \circ g : B \to D &\implies (h \circ g) \circ f : A \to D \\
\text{Let } x \in (g \circ f) \circ h & \qquad \text{Let } x \in (h \circ g) \circ f \\
\iff g(f(x)) \circ h \text{ and } h(g(x)) \circ f & \qquad \text{by def'n of functions} \\
\therefore h(g(f(x))) = h(g(f(x))) & \qquad \text{by def'n of functions } \square
\end{aligned}$$

1.2.4 Inverse Functions

A bijective function $f : A \to B$ that means that there exists a well-defined function $g : B \to A$ which undoes the effect of f. Now their compositions $g \circ f : A \to A = i_A$ and $f \circ g : B \to B = i_B$. The function g which implies these properties is known as the inverse function of f denoted by f^{-1}. It is well-defined for bijective functions itself being bijective.
Consider two functions f and g

$$f : A \to B \qquad\qquad g : C \to D$$

This implies that $g \circ f : A \to D$ if it is a bijective function has an inverse $f \circ g : D \to A$.

1.2.5 Character Persistence in Composition

If $f : A \to B$ and $g : B \to C$ are injective functions show that $g \circ f$ is an injective function also.

$$\begin{aligned}
&\text{Consider } g \circ f : A \to C && \text{by definition of composition}\\
&\text{Let } x_1, x_2 \in A : g \circ f(x_1) = g \circ f(x_2) &&\\
&\iff g(f(x_1)) = g(f(x_2)) && \text{by definition of composition}\\
&\text{Let } y_1 = f(x_1), y_2 = f(x_2) &&\\
&\implies g(y_1) = g(y_2) &&\\
&\iff y_1 = y_2 && \text{by the injectivity of g}\\
&\implies f(x_1) = f(x_2) &&\\
&\iff x_1 = x_2 && \text{by the injectivity of f}
\end{aligned}$$

$\therefore g \circ f$ is an injective function $\square$

The converse of this proof would be that if a composite function is injective then the functions comprising it must be injective also. A counter example to this would be the function $f \circ g(x) = e^{2x}$ which is the composite function of $f(x) = x^2$ a non-injective function and $g(x) = e^x$ an injective function.

If $f : A \to B$ and $g : B \to C$ are surjective functions show that $g \circ f$ is a surjective function also.

$$\begin{aligned}
&\text{Consider } g \circ f : A \to C && \text{by definition of composition}\\
&\text{Let } z \in C &&\\
&\iff y \in B, \exists z : g(y) = z && \text{by surjectivity of g}\\
&\iff x \in A, \exists y : f(y) = x && \text{by surjectivity of f}\\
&\implies x \in A, \exists z \ : g(f(y)) = z &&\\
&\iff x \in A, \exists z \ : g \circ f(x) = z && \text{by definition of composition}
\end{aligned}$$

$\therefore g \circ f$ is surjective also $\square$

1.2.6 A function of Symmetry !

Definition If S is a nonempty set then S_A is the set of all the *one-to-one* functions of S *onto* itself. In other words, the set of all bijective functions of S onto itself S
The following properties for the set S_A arise : $\sigma, \tau, \mu \in A_s$

1. $\exists \sigma \circ \tau \in S_A$ closure
2. $\mu \circ (\sigma \circ \tau) = (\mu \circ \sigma) \circ \tau$ associativity
3. $\exists i \in A_s : \sigma \circ i = i \circ \sigma = \sigma$ an identity element
4. $\exists \sigma^{-1} \in S_A : \sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = i$ an inverse element

Remark : Composition is not necessarily commutative within such a set.

Chapter 2

Groups

With the preceding notion of the set of bijections of a set onto itself we are in a good shape to abstract or generalise this notion to define the mathematical structure of Groups. This will become much clearer when *Cayley's Theorem* is considered later on.

Definition A nonempty set of elements G is said to form a group if in G there is defined a binary operation, called the product and denoted by $\cdot$ and thus the group can be represented as a pair $(G, \cdot)$ and must have the following characteristics.

1. $a, b \in G \implies a \cdot b \in G$ closure
2. $a, b, c \in G \implies a \cdot (b \cdot c) = (a \cdot b) \cdot c$ associativity
3. $\exists e \in G : a \cdot e = e \cdot a = a \forall a \in G$ identity element
4. $\forall a \in G \exists a^{-1} \in G : a \cdot a^{-1} = a^{-1} \cdot a = e$ inverse element

2.0.1 Preliminary Lemmas

With this algebraic structure introduced it is now about time to say some things about it.

The Identity element of a Group is Unique

Proof. Given that this is a uniqueness proof this will follow by assuming that there exist two identity elements and showing that these two such elements must be equal to one another.

Suppose the group G has two identity elements

$$\begin{aligned}
&\implies \exists\, e \in G : a \cdot e = e \cdot a = a \;\forall a \in G && \text{def'n of identity element}\\
&\text{Also } \exists\, f \in G : a \cdot f = f \cdot a = a \;\forall a \in G && \text{def'n of identity element}\\
&\implies e \cdot a = f \cdot a\\
&\implies e \cdot aa^{-1} = f \cdot aa^{-1} && \text{existence of inverse in } G\\
&\implies e = f && \text{as required.}
\end{aligned}$$

□

The Inverse element of an element in a Group is Unique

Proof. This is a uniqueness proof and so the basis of it must be two start with two things and show that they are in fact the same thing. The starting point will be to prove the cancellation law which will the lead to the desired uniqueness proof.

Lemma: Let $a \cdot x = a \cdot y \,\forall a, x, y \in G$

$$\begin{aligned}
&a^{-1}a \cdot x = a^{-1}a \cdot y && \text{existence of inverse in } G\\
&\implies e \cdot x = e \cdot y && \text{by def'n of inverse}\\
&\implies x = y
\end{aligned}$$

Now, suppose there exists $g \in G$ which has two inverses

$$\begin{aligned}
&\implies g \cdot g^{-1} = g^{-1} \cdot g = e && g \cdot g^{-2} = g^{-2} \cdot g = e\\
&\implies g \cdot g^{-1} = g \cdot g^{-2}\\
&\therefore g^{-1} = g^{-2} && \text{By the cancellation law}
\end{aligned}$$

□

The inverse of the inverse of an element is the element

Proof. The proof shall follow making use of the results achieved above that a group has a unique identity and inverse element for every element within it.

$$\begin{aligned}
&\text{Consider } g^{-1} \cdot \left(g^{-1}\right)^{-1} = \left(g^{-1}\right)^{-1} \cdot g^{-1} = e \;\forall\; g \in G && \text{inverse - } G \text{ is a group}\\
&\implies gg^{-1}\left(g^{-1}\right)^{-1} = ge\\
&\left(g^{-1}\right)^{-1} = ge\\
&\left(g^{-1}\right)^{-1} = g
\end{aligned}$$

□

Form of the inverse of a product

Proof.

$$\begin{aligned} (ab)^{-1} \cdot (ab) = e \ \forall a, b \in G & \qquad \text{inverse - } G \text{ is a group} \\ (ab)^{-1} \cdot (ab)b^{-1} = eb^{-1} & \\ (ab)^{-1} \cdot a = eb^{-1} & \\ (ab)^{-1} \cdot aa^{-1} = eb^{-1}a^{-1} & \\ \therefore (ab)^{-1} = b^{-1}a^{-1} & \end{aligned}$$

□

2.1 Notable Groups

2.1.1 Formalisms

Abelian Groups are groups whose binary operation is **commutative** for every element in the group, say G

$$\forall a, b \in G, \ a \cdot b = b \cdot a$$

It follows that non-abelian groups are groups which are non-commutative.
An example of an abelian Group is the group of integers under addition $(\mathbb{Z}, +)$ whilst an example of a non-abelian is that of 2×2 matrices under outer product matrix multiplication $\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \circ \right)$.

The Order of a Group shows the number of elements within a Group and is pertinent for *finite groups*. This notion is denoted $\circ(G)$. For dihedral groups the order is halved such that for example D_4 has 8 elements in total.

The Symmetric Group is the group of all bijective functions of a set ,S, with n elements,on-to itself. The group has $n!$ elements and represents the permutations of S. A permutation group is a subgroup of the symmetric group.

Generating Set of a Group is the set of group elements such that self application of these elements would generate all the elements of the group.
In practice generators are used primarily within *group presentation* together with a condition to full describe a group, as an example finite **cyclic groups** can be presented as

$$< a | a^n = e >$$

The finite **dihedral group** is given by two elements in its generating set leading to the presentation

$$< a, b | a^n = e = b^2, a^{-1}b = ba >$$

Cayley Tables Another way of representing groups which is perhaps analytical in that it exposes several characteristics of a group is the Cayley table named after the man that introduced them Arthur Cayley.These tables are also sometimes referred to as multiplication tables as they show the products of all the elements of a finite group. Above the cyclic group was presented as $< a|a^n = e >$ considering the cyclic group of order 4 the following Cayley table can be constructed.

C_4	e	a	a^2	a^3
e	e	a	a^2	a^3
a	a	a^2	a^3	e
a^2	a^2	a^3	e	a
a^3	a^3	e	a	a^2

It is evident that the table is symmetric across the main diagonal and this is evidence that the cyclic group is Abelian.
It is also evident that each row or column of the table (barring the outer column and top row) feature no repeated elements forming a latin square and displaying the permutative nature of groups and the fact that the cancellation law holds for groups given that a product cannot be formed in more than one distinct way. Cayley tables are also able to display ismorphism, consider the modulo multiplication group of order 5.Remember that in modular arithmetic $3 \times 2 = 6 = 1 \mod 5$ within the spirit of eliminating all the packets of 5 and leaving the remainder in this case, 1 packet of 5 and a 1 left over. That 1 is known as the residue.This will be elaborated over in detail in sections to come.

$x \mod 5$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

As can be seen when comparing the two Cayley tables of the C_4 group and the M_5 group they share an identical structure in *disguise* and this is known as an ismorphism between two groups denoted $C_4 \simeq M_5$.Similarly also the set [5] under modulo addition is also isomorphic to C_4.
More on group relations will be discussed in sections to come.

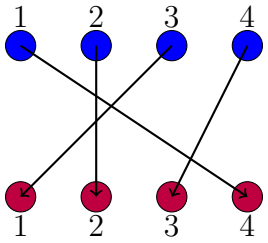
Finally, ismorphisms become dually more evident when expressing each row of a Cayley table using permutation cycles which are explained in the following section

$C_4 \simeq M_5$		
e	(1)(2)(3)(4)	1st row
a	(2431)	2nd row
a^2	(14)(23)	3rd row
a^3	(1342)	4th row

2.1.2 Permutation Groups

As mentioned previously, the symmetric group is the group of all bijective functions of a set on-to itself and a permutation is a subgroup of this group.Meaning that it represents only a number of these bijective functions and not all of them.

Notation Considering the set [4] an example of a permutation on it,ϕ might be



This can be denoted

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

Where the top row represents the domain and the bottom row the range.

Cycles and Orbits Another perhaps more telling representation is the following

$$(143)(2)$$

Which is read *one goes to four which goes to three which goes back to one and two goes back to two.* The orbit of 1 is the set of elements which 1 switches with in this cases 1,3 and 4.These two pairs of brackets represent two distinct cycles which together represent the initial permutation.It is in fact given as a proof on page 78 of Herstein that *every permutation may be represented as a product of disjoint cycles* where if one has two cycles , A and B, their product, AB, would be first carrying out the cycle B and carrying out on the image of B, A.
A cycle may be split into a product of 2-cycles which are also known as *transpositions* and as a result *every permutation can be decomposed to a product of 2-cycles.*

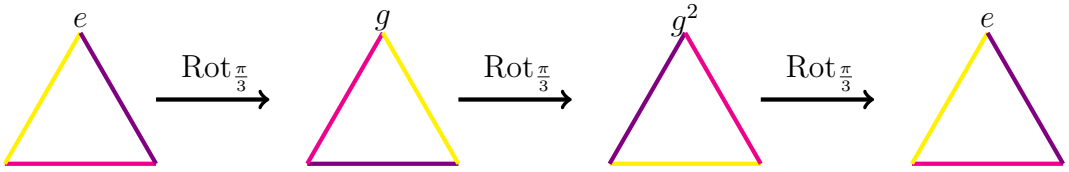
$$(14)(43)(31)$$

This is an odd permutation given that it is composed of an odd number of 2-cycles, even permutations are defined similarly.
An identity can only be represented by an even permutation ex: (56)(65).
With this in mind splitting the set of all permutations into two categories we have the odd permutations and the even permutation and for the reason preceeding this statement the odd permutations do not form a sub-group given that the identity element is an even permutation.All the even groups form a normal subgroup of the symmetric group known as an **alternating group**.

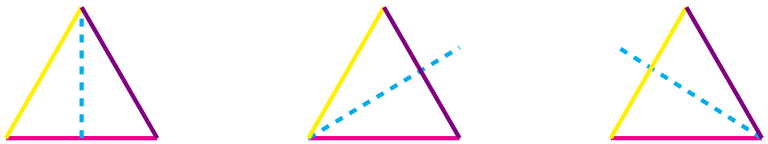
2.1.3 Visualising Groups through Symmetries

Cyclic Group As introduced above, a cyclic group is one which has a single element in its generating set and whose identity element is equal to an element with an index equal to its order.In more plain English this would mean that an equilateral triangle rotated three times through 120° would give the initial arrangement and that a set of 3 vertices and a rotation of 120° form a cyclic group of order 3.

$$< g | g^3 = e >=< \text{Rot}_{\frac{\pi}{3}} >$$



Dihedral Group Earlier it was discussed that a dihedral group is generated by two elements, one of which takes the form of the cyclic group generator which when having an index equal to the order of the group must be equal to some other generator applied twice. Ok, great. This means that knowing what was learned for a cyclic group we have rotations going on and another operation which bring us back to the initial state when carried out twice, a reflection!The number of vertices is related to the number of possible reflections.Thus, the group is generated by n reflections and n rotations having $2n$ elements in total.
Note that the reflection operation can only occur in one way $\because b^2 = e$, so different reflections are a direct result of carrying out a reflection in conjunction to a rotation.



To make what has just been described more evident the Cayley table for D_3 takes the form

D_3	e	a	a^2	b	ba	ba^2
e	e	a	a^2	b	ba	ba^2
a	a	a^2	e	ba	ba^2	b
a^2	a^2	e	a	ba^2	b	ba
b	b	ba	ba^2	e	a	a^2
ba	ba	ba^2	b	a	a^2	e
ba^2	ba^2	b	ba	a^2	e	a

As a final remark, the Dihedral group is non-Abelian for $n > 2$ which is evidenced by the generator relation $a^{-1}b = ba$

2.2 Subgroups

In the preceeding section were Groups were introduced it was seen that Groups are pairs of sets and a binary operation.Being a set, it would make sense that it has subsets and in this section our main point of inquiry will be a special form of these subsets which also form a group under the same binary operation of the group this set is found in, its *parent* group. This notion of a *group within a group* is referred to by the term **subgroup**.

Definition A nonempty subset H of a group G is said to be a subgroup of G if, under the product in G, H itself forms a group. This property is transitive meaning if H is a subgroup of some group G and K is a subgroup of H then K is a subgroup of G also. Subgroups may be denoted $H \leq G$.

2.2.1 A Criterion for Subgroups

What's the minimum requirement a subset of Group needs to meet to be considered a subgroup. Must it satisfy the 4 characteristics which define a group ? The answer is no !
It turns out that showing that closure and the existence of an inverse hold is equivalent to showing that a subgroup possesses the two other group characteristics of the existence of the identity element and associativity.

Lemma 2.2.1. *A nonempty subset H of the group G is a subgroup of G*

1. $\iff a, b \in H \implies ab \in H$ *closure*
2. $\iff a \in H \implies a^{-1} \in H$ *existence of an inverse*

Proof. First ($\implies$) assuming that H is a subgroup of G then quite trivially H is a group under the same binary operation under which G is a group by definition of the presumed subgroup quality. Given that H is a group then it must be closed (1) and it must possess inverse elements for each of its elements (2).
Now,($\impliedby$) conversely if H is a non-empty subset of G and possesses (1) and (2) it will be shown that it is a group possessing all 4 characteristics which define a group.

Associativity holds in H given that it is the subset of G where associativity holds.Now for the existence of the identity element

$$\begin{aligned} \text{Let } a \in H &\implies a^{-1} \in H && \text{by (2) existence of an inverse element} \\ &\implies aa^{-1} \in H && \text{By (1) closure in } H \\ &\implies e \in H && \text{By definition of } e \end{aligned}$$

□

This proof becomes prettier for a finite group as this can replace the existence of an inverse element as a requirement for a subset to be a subgroup, together ,again, with closure.

Lemma 2.2.2. ***-An Alternate Criterion***
If H is a nonempty finite subset of a group G and H is closed under the binary operation of its parent group, then H is a subgroup.

Proof. The meaning of this statement is that; to check whether a finite subset of a group is a subgroup all that is needed to be shown is whether or not this subset is **closed** under the binary operation of the parent.
To prove this we show that a nonempty finite subset possessing closure is a subgroup by evidencing that the two criterions of finiteness and closure give the characteristic of the existence of the inverse.

Let H be a finite subset of a group, G

$$\begin{aligned} &\text{Consider the trivial case } H = \{e\} \\ \implies\ & ee^{-1} = e^{-1}e = e && \text{Inverse holds for trivial case} \end{aligned}$$

Now for the case where $H \neq \{e\}$

$$\implies\ h \in H, hh \in H, ..., h^m \in H.... \in H \qquad \text{Closure of } H$$

For H to be a finite subset there must be repetitions in G

$$\begin{aligned} &\exists r, s \in \mathbb{Z} : r > s > 0 \wedge h^r = h^s \in G \\ & h^{r-s} = e && \text{By the cancellation law in } G \\ & s - r > 1 && \text{since otherwise } h = e \\ \implies\ & hh^{s-r-1} = e && \text{evidently, ghodd} \\ \implies\ & h^{s-r-1} = h^{-1} && \text{by definition of } e \\ \therefore\ & h^{-1} \in H && \text{Inverse exsits in } H\ ! \end{aligned}$$

H is shown to be a subgroup if it is a finite subset possessing closure

□

2.2.2 Some Examples of Subgroups

Subgroups of the Integers under Addition Consider $(\mathbb{Z}, +)$ is a a group, say G. Firstly, G is itself a group because $\mathbb{Z}$ is closed under addition, addition is associative, an additive inverse exists for each element also an identity element-0 exists in the set,

thus $(\mathbb{Z}, +)$ is a group.Now what are the subgroups of this group ?
Surely we must think of subsets and Herstein points us in the direction of the multiples of 5. Thinking about this by **Lemma 2.2.1** it is known that all that is needed is the existence of an inverse and closure for a subset to be a group. Closure is clearly inherited from the parent group and the existence of an inverse is given the product $5\mathbb{Z}^-$ which is in the subgroup, thus $5\mathbb{Z}$ is a subgroup.

Can this notion be extended to multiples of any integer ? Or perhaps more staunchly
Proposition 2.2.3. *All subgroups $H_n \leq \mathbb{Z}$ are of the form $n\mathbb{Z}$*

Proof. Proceeding by contradiction.
Suppose $\exists\, W \leq \mathbb{Z} : W \neq n\mathbb{Z}, \forall n \in \mathbb{Z} \wedge W \neq \{0\}$ and consider the smallest positive integer W say h which is obtainable by the well-ordered nature of $\mathbb{Z}$.
This h can form a subgroup within the subgroup W of the form $h\mathbb{Z}$ but under the supposition there should exist some $y \in W \wedge \notin h\mathbb{Z}$.
Consider the $\gcd(y, h)$ of these two numbers in W. By definition it can be expressed

$$\gcd(y, h) = \lambda y + \mu h \;\exists\; \lambda, \mu \in \mathbb{Z}$$

$$\implies \gcd(y, h) < y \wedge \gcd(y, h) < h \wedge \gcd(y, h) < 0 \;\text{※}$$

By definition the $\gcd(y, h)$ is smaller than y and h and it is also larger than 0 given that both are positive integers. But, this cannot be ! Since being smaller than h implies that $\gcd(y, h)$ and thus the contradiction arises proving that subgroups in $\mathbb{Z}$ under addition are of the form $k\mathbb{Z}$. □

Cyclic Subgroups Consider the notion that any group, say G, can be viewed as the disjoint union of its *singleton* subsets

$$G = \bigcup_{g \in G} \{g\}$$

Now consider that that the singleton subset for g is contained in the subgroup generated by G

$$G = \bigcup_{g \in G} \{g\} \leq \bigcup_{g \in G} \langle g \rangle \leq G$$

$$\implies G = \bigcup_{g \in G} \langle g \rangle$$

Where each of the subgroups are cyclic.

Modulo Congruence As introduced in **section 1.1.5**, modular arithmetic is an equivalence relation but this fact can be restated using subgroups by showing that reflexivity, symmetry and transitivity hold using the properties. So let's redefine modulo congruence in terms of groups

Definition 2.2.1. Let G be a group with $H \leq G$. For $a, b \in G$ we say that a is **congruent** to b as $a \equiv b \mod H$ if $ab^{-1} \in H$
Lemma 2.2.4. *The relation $a \equiv b \mod H$ is an equivalence relation.*

First let's make sense of the definition given in 2.2.4 and recall the example from page 6 where the notion of modular arithmetic was introduced using $n|(a-b)$.
Example: $6 \mod 5 = 1$ and 6 is *congruent* to 1 *iff* $1(-6) \in [5] \implies -5 \in 5\mathbb{Z}$ thus the congruence is satisfied.What's happening here in a coarse sense is we're collecting the packets of H and removing them from within the number b to achieve a meaning that *the difference of a and b must be perfectly divisible by n* as given by the first definition of modulo congruence in these notes *iff* $n|(a-b)$.

What does this mean in a group-y sense ?
This is what **Definition 2.2.4** achieves. Keeping with the example used above the over arching group G is in this case $\mathbb{Z}$ and a and b are both found in this group as 1 and 6 respectively. The definition also tells us that there is a subgroup H which is the basis of the *congruence* relation such that if the group product of a with the inverse element of b exists in H then a and $b \mod H$ are congruent.

Let's show the equivalence of these two definitions of additive **modulo congruence**!

Proof.

For reflexivity $a = a \mod H$ is shown:

$$\begin{aligned} \text{To be shown } aa^{-1} &\in H && \text{by } \textbf{Definition 2.2.4} \\ aa^{-1} &= e && \text{by def'n of identity element} \\ e &\in H && H \text{ is a subgroup} \\ \implies aa^{-1} &\in H \\ \therefore a &= a \mod H \end{aligned}$$

For symmetry $a = b \mod H \implies b = a \mod H$

$$\begin{aligned} \text{To be shown } ab^{-1} \in H &\implies ba^{-1} \in H && \text{by } \textbf{Definition 2.2.4} \\ \text{Consider } ab^{-1} &\in H && \text{premise.} \\ ab^{-1^{-1}} &\in H && \text{By existence of inverse in subgroups} \\ \left(b^{-1}\right)^{-1} a^{-1} &\in H && \text{By 2.0.1 Preliminary Lemmas} \\ \implies ba^{-1} &\in H && \text{By definition of inverse} \\ \therefore a = b \mod H &\implies b = a \mod H \end{aligned}$$

For transitivity $a = b \mod H \wedge b = c \mod H \implies a = c \mod H$

$$\begin{aligned} \text{To be shown } ab^{-1} \in H \wedge bc^{-1} \in H &\implies ac^{-1} \in H && \text{by } \textbf{Definition 2.2.4} \\ &\implies (ab^{-1})(bc^{-1}) \in H && \text{by def'n of subgroup} \\ &\implies aec^{-1} \in H && \text{by def'n of identity element} \\ &\implies ac^{-1} \in H \end{aligned}$$

$\therefore a = b \mod H \wedge b = c \mod H \implies a = c \mod H$

□

In doing this the notion of modulo congruence has been generalised to satisfy any operation and indeed any group. Whilst we were familiar with it for the group of integers under addition, it has now been shown that it is a concept which exceeds this boundary.

2.2.3 Cosets

Above we showed that *congruence* is an equivalence relation. Extending this notion leads to a powerful substructure of groups known as **cosets**.

As seen on page 22, congruence was introduced as *the group product of a with the inverse element of b in some subgroup, H*. If this is satisfied then congruence is attained.But, by nature of products we can multiply from either **left** or **right** to achieve a product leading to two distinct equivalence classes, the **left and right cosets**.
Recall that equivalence classes paritition !
Let's introduce cosets more formally.
Definition 2.2.2. Given a subgroup, H, we define two relations $\sim_L$ and $\sim_R$ on G

$$g_1 \sim_L g_2 \iff g_1^{-1} g_2 \in H \qquad\qquad g_1 \sim_R g_2 \iff g_2 g_1^{-1} \in H$$

This binary relation is evidently an equivalence relation as a result of **Lemma 2.2.4.** as such, the equivalence classes associated with this relation are known as the **left and right cosets** denoted $[g_2]_L$ and $[g_2]_R$, respectively.
Proposition 2.2.5. *Any left coset of the subgroup H of G has the form $gH = \{gx : x \in H\}$, while any right coset has the form $Hg = \{xg : x \in H\}$.*

Proof. Consider the right cosets as what will be said will follow symmetrically for the left cosets.
Now, in the words of my old philosophy teacher, "(we) start from what (we) know" which at the moment is the definition of cosets as equivalence classes. Thus the claim is restated as

$$[g]_R = Hg$$

meaning we will have to show both the forward and backwards argument by a subset relation.

($\implies$) Going forwards to prove $[g]_R \subseteq Hg$

Let $x \in [g]_R$	
$x \sim_R g$	by def'n of equivalence class
gx^{-1}	by def'n of $[g]_R$
$gx^{-1} \in H$	by closure
$(gx^{-1})^{-1} \in H$	by existence of inverse
$xg^{-1} = h \; \exists \; h \in H$	by def'n of inverse

Now post-multiplying by g

$xg^{-1}g = hg$	
$xe = x = hg$	by def'n of identity
$\implies x = hg \in Hg$	The forward implication is proven !

($\impliedby$) Going backwards to prove $Hg \subseteq [g]_R$

Let $x \in Hg$	
$\implies x = hg \; \exists \; h \in H$	
$x^{-1} = (hg)^{-1}$	taking the inverse
$x^{-1} = g^{-1}h^{-1}$	by def'n of inverse

Pre-multiplying by g

$gx^{-1} = g^{-1}gh^{-1} = eh^{-1} = h^{-1}$	by def'n of identity
$gx^{-1} \in [g]_R \wedge h^{-1} \in H$	by def'n of right coset and subgroup
$\implies h^{-1} \in [g]_R$	
For affermation $h = eh^{-1}$	by def'n of identity
$h^{-1} = gg^{-1}h^{-1}$	by def'n of identity
$g(hg)^{-1} \implies hg \in [g]_R$	The backwards case is definitely true.

$$\therefore [g]_R = Hg$$

□

Above we showed that $[g]_R = Hg$ and this Hg is the equivalence class of g in G and by **Theorem 1.1.1.** on page 6 it was shown that equivalence classes partition G into disjoint subsets.
This also means that any two right cosets of H in G are either identical or have no element in common.

Visualising Cosets Recalling our friend the Dihedral group of Order 3 we have

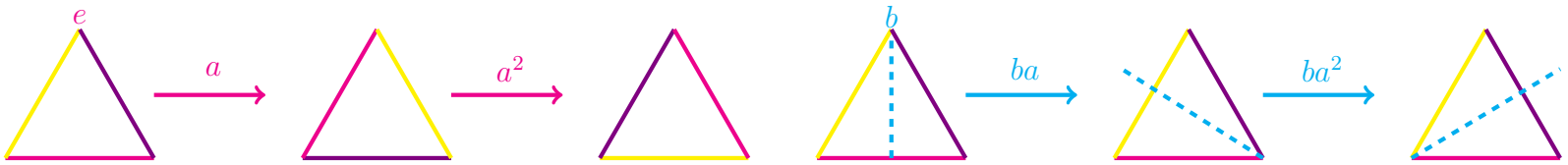
$$D_3 = \{e, a, a^2, b, ba, ba^2\}$$

Now we know that the cyclic group of order 3 is a subgroup of the dihedral group, let's call this H.
With what we established above we know that having this subgroup will allow us to partition the group into cosets, let's focus on the right cosets.

$[e]_r = He$ — Cyclic Rotations

$[b]_r = Hb$ — Reflections with Rotations

$$D_3 = \{e, a, a^2, b, ba, ba^2\}$$



The group has clearly been partitioned in to two cosets of apparently equal size, but is this always the case ? This is what will be examined in the section on Lagrange's Theorem.

Theorem 2.2.6. *Let H be a subgroup of G. Then there is a bijection between the left cosets and the right cosets of H in G.There are equally many of each.*

Proof. Consider that for any set of elements X in G, X^{-1} can be defined

$$X^{-1} = \{x^{-1} : x \in X\}$$

and evidently by nature of the inverse such a function is bijective.
Allow $X = Hg = \{hg : h \in H\}$

$$\implies X^{-1} = \{(hg)^{-1} : h \in H\}$$

which is viable since H being a subgroup, contains the inverse element

$$\implies X^{-1} = \{g^{-1}h^{-1} : h \in H\} = g^{-1}H^{-1} = gH$$

For affirmation $H = H^{-1}$ is evidently true given that a subgroup is a group and so contains inverse elements for each element in it, with this in mind the group can be viewed as its own inverse.
Thus having shown that the inverse of the right cosets are the left cosets then a function of the form

$$f : Hg \to gH$$

is a bijective function. □

2.2.4 Lagrange's Theorem

Pigeon-hole Principle describes the functional relationship between sets of different sizes, illustrating it through the use of pigeons flying into boxes, for some odd reason. It states *if $m > n$ and m pigeons fly into n pigeon holes then some pigeon hole has more than 1 pigeon.*
This fact can be equivalently stated in a number of ways which were eluded to in the preliminary section on functions but will be recounted here to recap.

1. If a function $f : A \to B$ is surjective then $|A| \geq |B|$ and is a many-to-one function such that $|A| \neq |B|$
2. It follows that if $f : A \to B$ has $|A| = |B|$ then such a function is bijective.

Lagrange's Theorem
Theorem 2.2.7. *There is a bijection between the right cosets of H in a finite group G which gives*

$$i \circ (H) = \circ(G)$$

Proof. Consider the function $f : h \to hg$ where $h \in H : H \geq G$. It is evidently well-defined $f(h) = hg$ by the well defined nature of the binary operation of G.

For injectivity, assume that $f(h_1) = f(h_2)$

$$\begin{aligned} f(h_1) &= f(h_2) && \text{premise} \\ h_1 g &= h_2 g && \text{by def'n of the function} \end{aligned}$$

Now post multiplying both sides by the inverse of g

$$\begin{aligned} h_1 g g^{-1} &= h_2 g g^{-1} \\ h_1 e &= h_2 e && \text{by def'n of inverse} \\ h_1 &= h_2 && \text{by def'n of identity} \end{aligned}$$

Injectivity has been shown and thus all that is left to show that the function is bijective is to show that it is surjective.

$$\begin{aligned} &\text{Let } z \in Hg \\ \implies &z = hg \ \exists \ h \in H \\ &\text{But } f(h) = hg && \text{by def'n of the function} \\ \implies &f(h) = z \end{aligned}$$

Thus all of Hg is exhausted and the image of the function is equal to the codomain showing surjectivity.

Therefore the function is surjective and so **bijective**, showing that all cosets are all of the same size !

$$\circ(H) = |H| = |Hg|$$

This means that a groups is **equipartitioned** by its cosets given that cosets are equivalence classes and thus cover the whole group.

He	Hg_1	Hg_2	$\cdots$			$\cdots$	Hg_i

Thus since all cosets are of the same size and there are finitely many, i, then the size of G is expressible as

$$\circ(G) = i \circ (H)$$

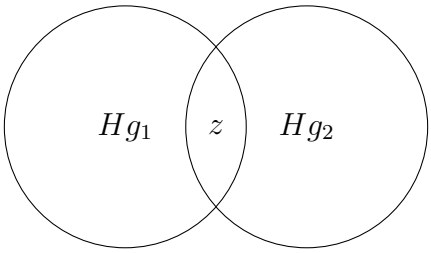
□

Definition 2.2.3. The **index** of H in G is the number of distinct right cosets of H in G. For finite groups, it is evident by Lagrange's theorem that

$$i = \frac{\circ(G)}{\circ(H)}.$$

Lagrange's Theorem if I didn't know what a Coset was. Within the first proof given, the definition of a coset as an **equivalence class** was heavily relied on in that the inheritance of the partitioning property was necessary.
Thus for a "full" or more exam friendly proof of Lagrange's theorem one proves the properties of cosets being **identical or disjoint** sets of **equal size** which **cover** the whole group for a more self contained version.

Proof. Firstly we consider the trivial case for which $H = \{e\}$.
Lagrange's theorem evidently holds as the subgroup order is 1 and one is a divisor of any other number.
Moving forward subgroups of the form $H \neq \{e\}$ together with their right cosets, Hg, will be considered. First the disjointness will be proven. To do this an element is assumed to be in the intersection of two cosets and it will be shown that as a result of this they are identical.



$$\begin{aligned} &\text{Let } z \in Hg_1 \cap Hg_2 && \text{premise} \\ \implies\ & z = h_1g_1 \ \wedge\ z = h_2g_2\ \exists\ h_1, h_2 \in H && \text{by def'n of intersection} \\ & h_1g_1 = h_2g_2 \end{aligned}$$

Premultiplying by h_1^{-1}

$$\begin{aligned} h_1^{-1}h_1g_1 &= h_1^{-1}h_2g_2 && \text{by def'n of inverse} \\ eg_1 &= h_1^{-1}h_2g_2 && \text{by def'n of identity element} \\ g_1 &= h_1h_2g_2 && \text{by existence of inverse} \\ &\because h_1h_2g_2 \in Hg_2 \\ \implies\ & hg_1 \in Hg_2\ \forall\ h \in H \\ &\therefore Hg_1 \subseteq Hg_2 \end{aligned}$$

By symmetry of the argument it follows

$$\begin{aligned} Hg_2 &\subseteq Hg_1 \\ \therefore Hg_1 &= Hg_2 \end{aligned}$$

Thus it has been shown that subsets are either identical or disjoint.
Now as before a function is constructed from the subgroup to its right cosets

$$f : h \to hg$$

and this is shown to be **well-defined**, **injective** and **surjective** giving a **bijection** as shown in the first proof given. That section would be repeated in this spot in the proof. Moving forward we prove that the cosets which at this point we have shown to be pairwise disjoint and of equal size, cover the whole group.We show that the union of pairwise disjoint cosets results in the group. Evidently this is an equality of sets and so we must show that both sets are subsets of eachother.

Starting with the forwards case ($\implies$)

$$\begin{aligned} &\text{Consider } Hg \in G && g \in G \ \wedge\ H \leq G, \text{ closure} \\ \implies\ & \bigcup_{g \in G} Hg \subseteq G && \text{subset proven.} \end{aligned}$$

For the backwards case ($\Longleftarrow$)

$$\begin{aligned} &\text{Let } x \in G \\ x = ex \in Hx & \qquad \text{since } e \in H \\ x \in \bigcup_{g \in G} Hg & \qquad \text{extending the argument} \\ \implies G \subseteq \bigcup_{g \in G} Hg & \qquad \text{since } x \in G \\ \therefore \bigcup_{g \in G} Hg = G & \end{aligned}$$

Thus the cosets are shown to be of the same size $\circ(H)$ and to partition G giving **Lagrange's Theorem**.

$$\circ(G) = i \circ (H)$$

□

Definition 2.2.4. The **order of an element** of a group or **period** is the least positive integer m such that

$$a^m = e$$

If no such integer exists, then a is said to have infinite order.
The order of an element is denoted $\circ(a)$.

With this definition in play we can now use it to hash out some corollaries of Lagrange's Theorem.

2.2.5 Corollaries of Lagrange's Theorem related to Elemental Order

Corollary 2.2.1. *If G is a finite group and $a \in G$, then $\circ(a) | \circ(G)$*

Proof. The corollary requires one to show that the elemental order is a divisor of the group order and thus it is natural, within the context, to think of a subgroup of order $\circ(a)$ so that Lagrange's Theorem can be applied.
This leads to the cyclic subgroup $\langle a | a^{\circ(a)} = e \rangle$ which will be shown to be of order $\circ(a)$.

Evidently, by definition of the cyclic group, the subgroup possesses at most $\circ(a)$ distinct elements but we also show that it cannot have less elements than this.
If this is the case then $a^i = a^j \ \exists \ 0 \leq i < j < \circ(a)$.

$$\implies a^{j-i} = e \text{ but } \nexists \ j, i : 0 < j - i < \circ(a)$$

$$\therefore \circ\langle a \rangle = \circ(a)$$

Given that the product of the identity element with the cyclic product is a coset, which as shown above would be of size $\circ(a)$, then applying Lagrange's Theorem

$$\circ(a) | \circ (G)$$

□

Corollary 2.2.2. *If G is a finite group $a \in G$, then $a^{\circ(G)} = e$.*

Proof.

$\circ(a) \| \circ (G)$	From the preceeding corollary
$\circ(G) = k \circ (a)$	by definition of *divisor*
$a^{\circ(G)} = a^{k \circ (a)}$	by equality
$a^{\circ(G)} = \left(a^{\circ(a)}\right)^k$	by laws of exponents
$a^{\circ(G)} = (e)^k$	by definition of order
$a^{\circ(G)} = e$	by definition of identity

□

Prime Orders for Cyclic Groups

Corollary 2.2.3. *If G is a finite group whose order is a prime number p, then G is a cyclic group.*

Definition 2.2.5. p is a prime number if and only if

1. $p|ab \implies p|a \vee p|b$
2. It has exactly two positive divisors, 1 and itself.

The proof will follow applying the definition of prime to Lagrange's Theorem.

Proof.

Recall Lagrange's Theorem $\circ(G) = k \circ (H)$

$$\implies \circ(G) = k \circ (H) = p : p|ab,\ p|a \vee p|b \qquad \text{by the premise}$$

Now evidently, k can take up two values depending on the size of H. k is either equal to p for the case that H is the singleton identity subgroup $\langle e \rangle$ or of value 1 for some subgroup of size p. Considering the second case

$$\implies \circ(G) = p \circ (e) \qquad \implies \circ(G) = \circ(H)$$

Considering the second case;

$$\circ(G) = \circ(H) = p$$

Suppose $\exists\, g : g \neq e, g \in G$ and let $H = \langle g \rangle$

$$\begin{aligned} \implies H = \langle g \rangle = \{e, g, g^2, ..., g^{p-1}\} \,:\, g^p = e \qquad & \text{By closure} \implies \circ(H) = p \\ \implies \circ(G) = \circ(H) = p & \\ \therefore G = H = \langle g | g^p = e \rangle \qquad & \text{as required.} \end{aligned}$$

□

Personally, I feel that this proof is note very clear but is along the lines of what Herstein and Prof.Sciriha present. I am of the stance that making use of **Corollary 2.2.1** and the notion of order would have made this proof much more evident in that it would be abundantly obvious that there can be no subgroup other than a cyclic subgroup in this context.
My reasoning would proceed as follows;

Proof.

$$\begin{aligned} \circ(a) | \circ(G) \qquad & \text{Corollary 2.2.1} \\ \circ(G) = k \circ (a) & \end{aligned}$$

For the non-trivial case, $k = 1$;

$$\circ(G) = \circ(a) = p$$

Now reflecting on the meaning of $\circ(a)$ (the period of an element such that it return to e after self application) leads one to the notion that $\exists$ some subgroup in G, say H, which must be of order p and **cyclic** by the def'n of $\circ(a)$

$$\begin{aligned} \implies \circ(G) = \circ(H) = p & \\ \therefore G = H = \langle g | g^p = e \rangle \qquad & \text{as required.} \end{aligned}$$

□

2.2.6 Corollaries of Lagrange's Theorem related to Number Theory

We look once more onto our friend modular arithmetic in this section perhaps more closely and within its own right.
In this section the notion of a congruence class is called upon once more such that the equivalence relation discussed will be that of $\times$ within the set $\mathbb{Z}$ which was introduced as modulo multiplication.

To the right is the Cayley Table for the modular group of order 5 in the integers. This was presented in previous sections but is given here to jog one's memory such that each group element represents the integer *left over* after all the packets of 5 have been taken away from the product.

$x \mod 5$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

Is it a Group though? Now in previous sections, what was most dealt with was modular addition whilst in this section we will be concerned with modular multiplication. The modular groups were not addressed with great care and detail but we're going to fix that. Firstly, performing a binary operation on a set of integers does not necessarily yield a group! To illustrate this let's apply $x \mod 6$ to the set of integers $[5]$ giving the following Cayley Table.

$x \mod 6$	1	2	3	4	5
1	1	2	3	4	5
2	2	4	0	2	4
3	3	0	3	0	3
4	4	2	0	4	2
5	5	4	3	2	1

Definition 2.2.6. The **Complete Residue System** of $a_i = i \mod m$ is the set of residues given by this relation for some $[m-1]$. Ex. The products in the Cayley Table give the complete residue system for $x \mod 6$.

The first thing which one notices on examining the Cayley Table is that it features 0s meaning that a product of two elements is not found in the set we started with, so we don't have **closure** in this set for $x \mod 6$.
The other missing thing for the set and binary operation to be considered a group; is an **inverse** element for each element in the set. Within this context for some set S we can define the inverse of the operation $x \mod n$ as

$$ab = aa^{-1} \equiv 1 \mod n \ : \ a, a^{-1}, b \in S$$

evidently only 1 and 5 have an inverse for this operation on this set, they also happen to be the only two elements whose products are all found within the set meaning that they have closure.
Pushing this idea forward it is almost obvious that dwindling this set down to just 1 and 5 would give a group under this operation, a **Residue Group** or reduced residue system.

$$R_6 = \{1, 5\}$$

R_6	1	5
1	1	5
5	5	1

Modular Multiplicative Inverse more closely In the example presented above it was shown that only elements which have a modulo multiplicative inverse can form a group, but what does it mean to have a modulo multiplicative inverse ?
Earlier the notion of a modular multiplicative inverse wass introduced in the following way

$$ab = aa^{-1} \equiv 1 \mod n \; : \; a, a^{-1}, b \in S$$

and referring back to the example given $5 \times 5 = 25$ which removing the *packets* of 6, ie: 24, leaves 1, the identity element. Similarly $1 \times 1 = 1$ which has no *packets* of 6 and so gives the identity element also. In this context both 5 and 1 are their own inverses.
Now we appreciate that both 1 and 5 share a particular property within the context of the number 6.

$$\gcd(1, 6) = 1 \qquad\qquad \gcd(5, 6) = 1$$

This means that 5 and 1 are both **Coprime** or relatively prime to 6! From this we are able to deduce that any integer possessing a modular multiplicative inverse is coprime to the modulus, in this case 6.

> For all $a \in [m - 1]$ coprime to m and so having the following property
>
> $$\gcd(a, m) = 1$$
>
> a modular multiplicative inverse, exists.

In general, it is through this relation that the Residue Group and Totient function are defined rigorously as will be done in some following section.

Another way to look at this is using the notion of **Zero Divisors**. These are non-zero elements in the set, the product of whom with some other non-zero element will result in a zero element. In other words such a product is equal to some multiple of the modulus and will have

$$ab = 0 \mod m \; : \; a, b \neq 0$$
$$\implies \gcd(a, m) \neq 1$$

Within the example given, $2 \times 3 = 6$ which is a *packet* of 6 leaving zero and so 2 and 3 are zero divisors.

Theorem 2.2.8. *A divisor of zero has no inverse*

Proof. Taking an approach by contradiction
Suppose $xy = 0 \; : \; x \neq 0, y \neq 0 \; \wedge x^{-1}$ exists

$$\begin{aligned} \implies xy &= 0 && \text{supposition.} \\ x^{-1}xy &= x^{-1}0 && \text{Pre-multiplying by } x^{-1} \\ ey &= 0 && \text{by def'n of identity and 0} \\ y &= 0 \; ※ \end{aligned}$$

□

Residue Groups and Euler's Totient function Having fully ironed out all the ideas needed to understand these concepts it is now time to formally introduce them and start to applying them to number theory.
Starting with the formalities;

Definition 2.2.7. Residue Group or reduced residue system, denoted R_n is a set of integers less than and coprime to a modulus m, under the binary operation $\times \bmod m$.
Theorem 2.2.9. *A reduced residue system forms a group in* $\mathbb{N}$ *under* $\times$ *mod* m

Proof. There are 4 Axioms that must be upheld for a set and operation pair to be thought of as a group. Closure, Associativity and the existence of the identity and inverse elements. For this theorem to be proven all four of these axioms must be shown to hold for $(\mathbb{Z}, \times \bmod m)$.

(i) Closure;
Let $r_1, r_2 \in R_n$. It will be shown that $r_1 r_2 \in R_n$ also for closure. To proceed it is recalled that if r_1 and r_2 are in a reduced residue class then they are both coprime to some modulus, n.

$$\gcd(r_1, n) = 1 \qquad \gcd(r_2, n) = 1 \qquad \text{by def'n of Coprime}$$

by Bézout's Lemma

$$\implies \gcd(r_1, n) = \lambda_1 r_1 + \mu_1 n = 1 \quad \gcd(r_2, n) = \lambda_2 r_2 + \mu_2 n = 1 \; \exists \; \lambda_1, \lambda_2, \mu_1 \, \mu_2 \in \mathbb{Z}$$

$$\implies \begin{cases} \lambda_1 r_1 + \mu_1 n = 1 \\ \lambda_2 r_2 + \mu_2 n = 1 \end{cases}$$

Multiplying these two equations out will allow for the formation of a linear combination of $r_1 r_2$ and n.

$$\lambda_1 \lambda_2 r_1 r_2 + \lambda_1 \mu_2 n + \lambda_2 \mu_1 n + \mu_1 \mu_2 n^2 = 1$$

$$\implies (\lambda_1 \lambda_2) r_1 r_2 + (\lambda_1 \mu_2 + \lambda_2 \mu_1 + \mu_1 \mu_2 n) n = 1$$

Letting $(\lambda_1 \lambda_2) = \alpha$ and $(\lambda_1 \mu_2 + \lambda_2 \mu_1 + \mu_1 \mu_2 n) = \beta$

$$\alpha r_1 r_2 + \beta n = 1$$

Thus reversing Bézout's Lemma

$$\therefore \gcd(r_1 r_2, n) = \alpha r_1 r_2 + \beta n = 1.$$

It is evident that $r_1 r_2$ is coprime to the modulus, n by definion of coprime and so is part of the Residue Group. Proving that Closure holds.

(ii) Associativity;
This property is inherited from $(\mathbb{Z}, \cdot)$

(iii) Identity Element;
1 can never be a zero-divisor and is coprime to any integer meaning that

$$1 \in R_n \, \forall \, n \in \mathbb{Z}$$

As a result an identity element exists in every Residue group.

(iv) Inverse Law holds;
This will follow by first showing what the structure for an inverse element in modular multiplication would look like and then proving that an inverse element of such a structure must be coprime to the modulus and thus be in the group.
Consider $\forall \, r \in R_n$

$$\gcd(r, n) = 1$$

by definition of the residue group with modulus n. Now applying Bézout's Lemma the gcd may be restated as

$$\lambda r + \mu n = 1 \; \exists \; \lambda, \in \mathbb{Z} \tag{2.1}$$
$$\lambda r = 1 - \mu n \tag{2.2}$$
$$\implies \lambda r = 1 \mod n \tag{2.3}$$

Now λ can be expressed as follows in terms of the modulus as a result of the previous line

$$\exists \; v, w \in \mathbb{Z} : \lambda = nv + w \text{ where } 0 \leq w \leq n - 1 \tag{2.4}$$
$$(nv + w)r = 1 \mod n \tag{2.5}$$
$$\implies wr = 1 \mod n \tag{2.6}$$
$$\implies r^{-1}r = 1 \mod n \tag{2.7}$$

Thus, structurally w takes the form of an inverse element. Now it shall be shown that such a form exists in the group. Proceeding by contradiction, assume $w \notin R_n$.

$$\implies \gcd(w, n) \neq 1 = \alpha \; \exists \; \alpha\mathbb{Z} \tag{2.8}$$
$$\implies \alpha | w \text{ and } \alpha | n \text{ by def'n of gcd.} \tag{2.9}$$
$$\implies \alpha | \lambda \tag{2.10}$$
$$\text{But, since } \lambda r + \mu n = 1 \text{ from (2.1)} \tag{2.11}$$
$$\implies \alpha | 1 \text{ and so } \alpha = 1 \; ※ \tag{2.12}$$

This contradicts the supposition that $a \neq 1$ and that w is not in the residue group. Proving that the Residue Group has an inverse and that ultimately, a reduced residue system forms a group.

$$r^{-1} \in R_n \tag{2.13}$$

□

Definition 2.2.8. Euler's Totient Function denoted $\phi(n)$ is the order or size of the residue group meaning that it indicates the number of integers less than and are coprime or relatively prime to, the modulus, n.

In the given example of R_6 the corresponding totient $\phi(6) = 2$ since $R_6 = \{1, 5\}$.

Corollary 2.2.4. ***Euler's Theorem*** *If n is a positive integer and it is relatively prime to n, then $a^{\phi(n)} \equiv 1 \mod n$.*

It is evident that this corollary is the application of **Corollary 2.2.2** from earlier to the Residue group such that $\phi(n) = \circ(G)$ and $e = 1 \mod n$ for $G = R_n$.
Prof.Sciriha carried out a proof which shows that **Corollary 2.2.2** may be adapted to the Residue group in such a way, making use of the division algorithm.

Proof. By division algorithm, consider that some group element of the residue group a, can be expressed as

$$a = qn + r \;\exists\; q, r \in \mathbb{Z} \;:\; 0 \leq r \leq n - 1.$$

That is, q packets of n and some other amount r. Given that the premise is about taking the order of the group as an exponent

$$\implies a^{\phi(n)} = (qn + r)^{\phi(n)}.$$

Now recall that $(a + b)^k = \sum_{i=0}^{k} \binom{k}{i} a^i b^{k-1}$.

$$\implies (qn + r)^{\phi(n)} = \sum_{i=0}^{\phi(n)} \binom{\phi(n)}{i} qn^i r^{\phi(n)-i}$$

Taking out all terms containing packets of n

$$a^{\phi(n)} = r^{\phi(n)} \mod n$$

Structurally this is what we want as when **Corollary 2.2.2** is applied $r^{\phi(n)} \mod n$ becomes $1 \mod n$ as required but before making this step r must be shown to be within the group. This will be shown employing the tactic used when proving that the inverse element lies in the group.

Suppose by contradiction that $r \notin R_n$

$$\implies \gcd(r, n) \neq 1 = \alpha \;\exists\; \alpha \in \mathbb{Z}$$

Which by definition of gcd would imply that

$$\alpha | r \wedge \alpha | n.$$

Now recalling that r is part of reformulation of a and Bézout's Lemma

$$a = qn + r \qquad\qquad \lambda a + \mu n = 1.$$

$$\implies \alpha|a \wedge \alpha|n$$

$$\therefore \alpha|1 \text{ ※}$$

This cannot be since $\alpha \neq 1$ by supposition of the contradiction, thus r is in the group and so the argument alluded to earlier will follow by the application of **Corollary 2.2.2**

$$\because r \in R_n \implies a^{\phi(n)} = r^{\phi(n)} \mod n$$

By **Corollary 2.2.2**

$$r^{\phi(n)} = e = 1 \mod n = a^{\phi(n)}$$

$$\therefore a^{\phi(n)} = 1 \mod n$$

□

Considering again our friend R_6 let's apply Euler's Theorem to 5

$$5^2 = 25 = 1 + 24 = 1 + 6(3) = 1 \mod 6$$

$$a^{\phi(n)} = 1 \mod 6$$

The theorem thus holds and 1 is clearly r.

Corollary 2.2.5. ***Fermat's Little Theorem*** *If p is a prime number and a is any integer, then $a^p = a \mod p$.*

This is clearly an application of Euler's theorem to prime numbers where any integer less than the prime number is relatively coprime to the prime number, also, by definition and as such forms a residue with p as the modulus of the group of size $\phi(p) = p - 1$.

Proof. This theorem follows from Euler's Theorem by nature and thus this is a natural starting point for the proof.
Let $a \in [p-1]$ where p is prime.

$$\implies \gcd(a, p) = 1 \text{ by def'n of prime}$$

It is evident as described previously, that the numbers less than p form a residue group with modulus p of size $\phi(p) = p - 1$. So by Euler's theorem

$$a^{p-1} = 1 \mod p$$

$$a^p \times a^{-1} = 1 \mod p$$

$$a^p = a \mod p \ \because 1 \mod p = 1 + wp \ \exists \ w \in \mathbb{Z}$$

This is sufficient for cases within the residue group and so less than p but what about numbers greater than p ? Integers that are larger than p are either coprime to p and so do not have p as a divisor or are coprime to p, for which the case reduces to the section of the proof given above.

Thus let us prove that Fermat's Little Theorem holds for integers larger than p that are not coprime to p. As mentioned earlier, such integers would have p as a divisor.

$$\gcd(a,p) \neq 1 \implies p|n$$

This would mean that

$$a = 0 \mod p.$$

Extending this argument

$$a^p = 0 \mod a \text{ also.}$$
$$\implies a^p - n = 0 \mod p \text{ which is equivalent to } p|a^p - a$$
$$\therefore a^p = a \mod p \text{ by def'n of mod.}$$

□

The presentation by Herstein is perhaps some what misleading given that Euler's Theorem is a specific case of Fermat's little theorem for when an integer is coprime to a prime number and so applicable to the residue group. But in truth Fermat's little theorem is about all integers and not just those less then some prime number.

Euler's theorem and Fermat's Little Theorem are central to the inner workings of the RSA, Rivest-Shamir-Adleman, cryptographic system and working out the prime factors of products of large primes.

2.3 Normal Subgroups

Initially pointed out by Galois, normal subgroups are ones which produce left and right cosets which are the same. This is the most intuitive definition but worry not, this shall be abstracted soon enough! To introduce this notion let's consider our old friend the Dihedral Group of Order 3 and the Symmetric Group of Order 3 as two non-separate examples.
The term non-seperate is indicative of the fact that $D_3 \simeq S_3$.

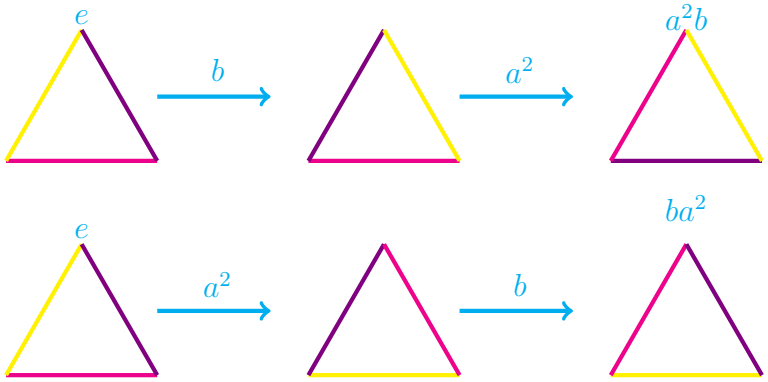
2.3.1 Visualising Normal Subgroups using D_3

$$D_3 = \langle a, b | a^3 = b^2 = e, a^{-1}b = ba \rangle$$
$$D_3 = \{e, a, a^2, b, ab, ab^2\}$$

First consider the subgroup $H = \{e, b\}$ and its left and right cosets;

Left Cosets	Right Cosets
$H = \{e, b\}$	$H = \{e, b\}$
$aH = \{a, ab\}$	$Ha = \{a, ba\}$
$a^2H = \{a^2, a^2b\}$	$Ha^2 = \{a^2, ba^2\}$

Given that a and b do not commute this means that the cosets are not equal, and this makes sense.
The subgroup $\{e, b\}$ corresponds to a reflection across an axis of symmetry of the triangle. The result of such a reflection is totally dependent on the starting orientation of the triangle and hence it is for this reason that the left and right cosets don't match up. This can be illustrated physically as shown below.



Alternatively, consider the cyclic subgroup C_3 which is embedded in the dihedral group, as was done in the section on visualising cosets, and let's consider the left and right cosets of this subgroup within D_3.

Left Cosets	Right Cosets
$C_3 = \{e, a, a^2\}$	$C_3 = \{e, a, a^2\}$
$bC_3 = \{b, ba, ba^2\}$	$C_3b = \{b, ab, a^2b\}$

At first one may be inclined to think that bC_3 and C_3b are not equal but recall the relation defining the dihedral group

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

Using this we show

$$\begin{aligned} ba^2 &= (ba)a \\ ba^2 &= a^{-1}(ba) \\ ba^2 &= a^{-1}(a^{-1}b) \end{aligned} \qquad \begin{aligned} & \\ ba &= a^{-1}b \\ ba &= a^{-1}b \end{aligned}$$

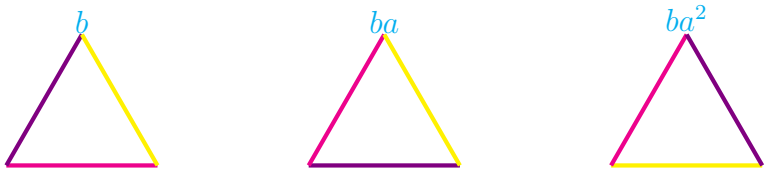
Consider the fact that the inverse of a within the group is a^2 such that $e = a^3 = aa^{-1} = aa^2$, giving

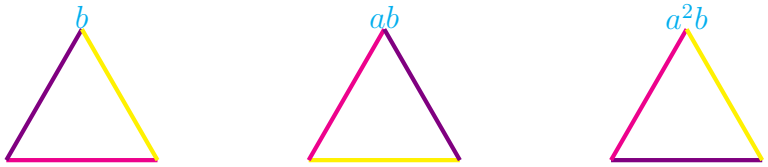
$$\begin{aligned} ba^2 &= a^4b \\ ba^2 &= a^3ab \\ ba^2 &= eab \\ \underline{ba^2 &= ab \in C_3b} \end{aligned} \qquad \begin{aligned} & \\ & \\ & \\ \underline{ba &= a^2b \in C_3b} \end{aligned}$$

Of course adding $b = b \in C_3b$ this would mean that all the cosets coincide !
As can now be appreciated whilst all the elements were there earlier, the distribution of

the elements meant that they were not within the same cosets. Now again considering this physically makes more sense.

Given that now we are considering the cyclic subgroup all the rotations are being carried out at every stage and reflected before hand or after, where as before it was only a singular rotation being carried out with a reflection. This means that given that all operations are being carried out, whilst they might not occur in the same order, they will occur within corresponding cosets.

$bC_3 =$ 

$C_3b =$ 

2.3.2 A Mathematical Definition of Normal Subgroups and Lemmas

Definition 2.3.1. A subgroup N of G is said to be a *normal subgroup* of G, denoted $N \trianglelefteq G$, if and only if $\forall\, g \in G$ and $n \in N$, $g^{-1}ng \in N$.
Equivalently this can be stated $N \trianglelefteq G \iff g^{-1}Ng \subseteq N \,\forall g \in G$ by definition of subset.

This mathematical definition seems disconnected from the notion of equality of left and right cosets presented earlier so we must demonstrate their interchangeability as definitions.
Before pursuing this interchangeability we look for a strengthening of the definition which will aid the proof.
Lemma 2.3.1. $N \trianglelefteq G \iff g^{-1}Ng = N$

Proof. For ($\implies$);

$$\begin{aligned} &\text{Let } N \trianglelefteq G && \\ &\forall\, g \in G, g^{-1}Ng \subseteq N && \text{by def'n of normal} \\ &\forall\, g \in G, gNg^{-1} \subseteq N && \text{by existence of inverse in } G \end{aligned}$$

Pre-multiplying by g^{-1} and post-multiplying by g

$$\begin{aligned} g^{-1}gNg^{-1}g &\subseteq g^{-1}Ng \\ eNe &\subseteq g^{-1}Ng && \text{by def'n of inverse} \\ \implies N &\subseteq g^{-1}Ng && \text{by def'n of identity} \\ \therefore g^{-1}Ng &= N \end{aligned}$$

For ($\impliedby$);

Consider $g^{-1}Ng = N$.
By definition of equality of sets, this readily gives $\forall\, g \in G$, $g^{-1}Ng \subseteq N$ and so that $N \trianglelefteq G$. □

Alternatively, in a more Prof.Sciriha like manner take an element in each of the sets and carry out the exact same procedure.

Lemma 2.3.2. *The subgroup N of G is a normal subgroup of G if and only if every left coset of N in G is a right coset of N in G.*

Making use of the definition introduced in this section let translate this into mathematics

$$N \trianglelefteq G \iff \forall g \in G,\, Ng = gN$$

Proof. For ($\implies$);

$$\begin{aligned} &\text{Let } N \trianglelefteq G \\ \forall\, g \in G, g^{-1}Ng &\subseteq N && \text{by def'n of normal} \end{aligned}$$

But using **Lemma 2.3.1** we can implement the stronger defintion

$$\forall\, g \in G, g^{-1}Ng = N \qquad \text{by } \textbf{Lemma 2.3.1}$$

Now, simply pre-multiplying by g

$$Ng = gN \qquad \text{as required.}$$

For ($\impliedby$);

$$\forall g \in G,\, Ng = gN$$

Pre-multiplying by g^{-1}

$$\begin{aligned} g^{-1}Ng &= N \\ \implies g^{-1}Ng &\subseteq N && \text{by def'n of set equality} \end{aligned}$$

□

Thus we have illustrated that the two definitions of a normal subgroup are interchangeable.

Proposition 2.3.3. *Let H be a subgroup of a group G. Each of the following conditions implies that H is a normal subgroup:*

1. *G is abelian*
2. *H is finite and the only subgroup of G of its order*
3. *H has index 2 in G*

For *(1)*.
Here we are saying that any subgroup of an Abelian group G, is normal. Firstly, thinking about this, this makes perfect sense right; normal subgroups are ones which produce left and right cosets which are identical and cosets are equivalence classes taking the group product as the equivalence relation, and if this product is commutative then it is no surprise that these cosets must be identical !

As an example consider the Cyclic group C_3 taking C_2 as a subgroup.

Left Cosets	Right Cosets
$eC_2 = \{e, a\}$	$C_2e = \{e, a\}$
$aC_2 = \{a, a^2\}$	$C_2a = \{a, a^2\}$

To prove this let us make use of the mathematical definition of normal subgroups directly.

Proof. Let $H \leq G : G$ is Abelian.

$$
\begin{aligned}
\text{Let } x \in g^{-1}Hg \in G & \qquad \text{possible by closure} \\
x = g^{-1}hg \; \exists h \in H & \\
x = g^{-1}gh & \qquad \text{Commutativity of Abelian Groups} \\
x = eh & \qquad \text{by def'n of inverse} \\
x = h \in H & \qquad \text{by def'n of identity} \\
\implies g^{-1}Hg \subseteq H &
\end{aligned}
$$

□

For *(2)*. Here we are saying that if there is some $H \leq G$ such that it is the **only** subgroup of G with order m then $H \trianglelefteq G$.

Proof. We begin by showing that there is some other subgroup in G which has a normal structure and by showing that it must have the same order as H we prove that it must indeed be H by the premise and so H has normal structure.

To show that some set is a subgroup it must have closure and an inverse for each of its elements by **Lemma 2.2.1.**

For Closure;
Let $x_1, x_2 \in g^{-1}Hg$

$$\begin{aligned}
\implies x_1 = g^{-1}h_1g,\ \exists\ h_1 \in H \qquad & x_2 = g^{-1}h_2g,\ \exists\ h_2 \in H \\
x_1x_2 = (g^{-1}h_1g)(g^{-1}h_2g) & \\
x_1x_2 = (g^{-1}h_1)(gg^{-1})(h_2g) & \\
x_1x_2 = (g^{-1}h_1)e(h_2g) & \text{By def'n of inverse} \\
x_1x_2 = g^{-1}(h_1h_2)g & \text{By def'n of identity} \\
\therefore\ x_1x_2 \in g^{-1}Hg & \text{By closure in } H
\end{aligned}$$

For Inverse;
Let $x \in g^{-1}Hg$

$$\begin{aligned}
\implies x = g^{-1}hg,\ \exists\ h \in H & \\
x^{-1} = (g^{-1}hg)^{-1} & \\
x^{-1} = (hg)^{-1}(g^{-1})^{-1} & \\
x^{-1} = g^{-1}h^{-1}g & \\
\therefore\ x^{-1} \in g^{-1}Hg \qquad & \text{By existence of inverse in } H
\end{aligned}$$

Therefore $g^{-1}Hg \leq G$.
Now recall that from the premise $\circ(H) = m : m \in \mathbb{Z}$.
Consider,

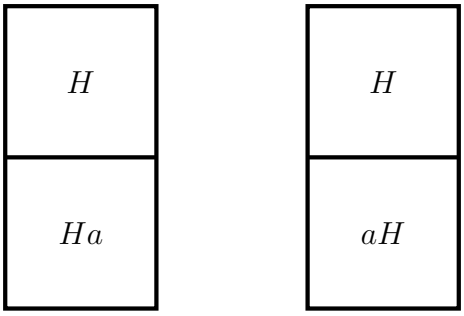
$$\begin{aligned}
h \cdot h \cdots h = h^m = e & \\
\underbrace{g^{-1}hg \cdot g^{-1}hg \cdots g^{-1}hg}_{m \text{ times}} = e & \\
\implies (g^{-1}hg)^m = e & \\
\therefore\ \circ(g^{-1}Hg) = m &
\end{aligned}$$

Given that by the premise; H is the **only** subgroup of order m in the group G then

$$H = g^{-1}Hg.$$

□

For (*3*). Here we are saying that any subgroup of index 2, ie: splits the group in half when forming cosets by Lagrange's Theorem, is a normal subgroup.



Proof. By Lagrange's Theorem

$$\circ(G) = i \circ (H).$$
$$\text{For this case } \circ(G) = 2 \circ (H).$$

This gives

Left Cosets	Right Cosets
$eH = H$	$He = H$
$gH = \{G \setminus H\}$	$Hg = \{G \setminus H\}$.

And so

$$Hg = gH.$$
$$\therefore H \trianglelefteq G.$$

□

Theorem 2.3.4. *A subgroup N of G is a normal subgroup if and only if the product of two right cosets gives a right coset*

$$N \trianglelefteq G \iff NaNb = Nab$$

Proof. For ($\implies$); We must show that normality infers a set equivalence thus we start by showing $NaNb \subseteq Nab$

$$\text{Let } x \in NaNb : a, b \in G \text{ and } N \trianglelefteq G$$
$$\implies x = n_1 a n_2 b \ \exists \ n_1, n_2 \in N$$
$$x = n_1 (a n_2) b \ \exists \ n_1, n_2 \in N$$

Now given that N is a normal subgroup and $n_2 \in N$;

$$an_2 = na_3 \ \exists \ n_3 \in N.$$
$$\implies x = n_1 n_3 ab$$
$$\implies x \in Nab \text{ by closure of } N \text{ as a subgroup}$$
$$\therefore NaNb \subseteq Nab \text{ by def'n of subset}$$

Now we show that $Nab \subseteq NaNb$

$$\text{Let } x \in Nab : a, b \in G$$
$$\implies x = n_4 ab \ \exists \ n_4 \in N$$
$$x = n_4 aeb \text{ by def'n of identity}$$
$$\implies x \in NaNb \text{ since } N \text{ is a subgroup}$$
$$\therefore Nab \subseteq NaNb \text{ by def'n of subset}$$
$$\implies N \trianglelefteq G \implies NaNb = Nab \text{ as required.}$$

For ($\Longleftarrow$);
Here we show that this equality infers normality.

$$\begin{gathered}\text{Given } NaNb = Nab \text{ let } x \in g^{-1}Ng \\ \implies x = g^{-1}n_5 g \ \exists \ n_5 \in N \\ x = eg^{-1}n_5 g \text{ by def'n of identity element}\end{gathered}$$

Now, given that $e \in N$ since N is a subgroup we have

$$x \in Ng^{-1}Ng.$$

Which from the premise

$$\begin{gathered}x \in Ng^{-1}Ng = Ng^{-1}g = Ne = N \\ \implies x \in N \\ \therefore g^{-1}Ng \subseteq N \text{ by def'n of subset.}\end{gathered}$$

□

Thus both sides of the bi-implication have been satisfied and a subgroup is normal if and only if the cosets it produces give products which are of the same parity of their constituents. **Preserving the structure** of parity of cosets.

This equation, whilst at first glance seemingly unremarkable unlocks a new level of abstraction within group theory through the introduction of **homomorphisms** and quotient structures. This is what will be examined in the coming section, this preservation of structure and how it can be abstracted from parity to any type of structure.

2.4 Homomorphisms and Quotient Groups

2.4.1 Quotient Structures

In **Proposition 2.3.3.** we split a group in half and then removed the normal subgroup that gave us the cosets, leaving only the cosets in the group. This is the process described by Prof.Irene as going from the villagers to the villages such that we are know representing types of elements by their cosets, and so people from Marsaskala by Marsaskala.
This set of cosets without the normal subgroup is known as the Quotient or factor structure and we claim that it forms a group under the relation

$$NaNb = Nab.$$

Definition 2.4.1. Let N be a normal subgroup in the finite group G. The Quotient group, $\frac{G}{N}$ is the set of (left or right) cosets of N in G with the binary operation

$$(Ng_1)(Ng_2) = Ng_1g_2.$$

Lemma 2.4.1. *$(Ng_1)(Ng_2) = Ng_1g_2$ is a binary relation.*

Proof. We show that the operation is well-defined.

Let $N \trianglelefteq G$ and $g_1, g_2, w_1, w_2 \in G$:

$$\begin{aligned} Ng_1 &= Nw_1 & Ng_2 &= Nw_2 \\ Ng_1Ng_2 &= Ng_1g_2 & Nw_1Nw_2 &= Nw_1w_2 \end{aligned}$$

Thus, to show well-definition, we wish to show that $Ng_1g_2 = Nw_1w_2$, an equality of cosets. Consider that

$$\begin{aligned} Ng_1 &= Nw_1 & Ng_2 &= Nw_2 \\ Ng_1w_1^{-1} &= N & Ng_2w_2^{-1} &= N. \end{aligned}$$

Recalling that $e \in N$ given that N is a subgroup then rewriting this out in terms of elements gives

$$\begin{aligned} eg_1w_1^{-1} = n_1,\ \exists\ n_1, e \in N & & eg_2w_2^{-1} = n_2,\ \exists\ n_2, e \\ g_1w_1^{-1} = n_1 & & g_2w_2^{-1} = n_2 \\ g_1 = n_1w_1 & & g_2 = n_2w_2 \end{aligned}$$

With these handy relations at our disposal let us show that the cosets are equal starting by showing that they share some element x.

$$\begin{gathered} \text{Let } x \in Ng_1g_2 \\ x = n_3g_1g_2,\ \exists\ n_1 \in N \end{gathered}$$

Using the pre-derrived relations

$$\begin{aligned} x &= n_3(n_1w_1)(n_2w_2) \\ x &= (n_3n_1)(w_1n_2)w_2) \end{aligned}$$

Given that N is normal

$$\begin{aligned} x &= (n_3n_1)(n_4w_1)(w_2) \\ x &= (n_3n_1n_4)(w_1w_2) \end{aligned}$$

By closure in N as a subgroup

$$\begin{gathered} x = n_5w_1w_2 \\ \therefore\ x \in Nw_1w_2 \end{gathered}$$

Now given that cosets are either equal or disjoint then as a result of $x \in Ng_1g_2$ and also $x \in Nw_1w_2$ we have

$$Ng_1g_2 = Nw_1w_2.$$

Thus showing that $(Ng_1)(Ng_2) = Ng_1g_2$ is a well-defined operation. □

Theorem 2.4.2. *Quotient structures indeed form quotient groups.*

Proof. We show that quotient structures as described above under the binary operation $NaNb = Nab$ indeed form a group satisfying the four axioms defining the algebraic structure of groups;

1. Closure
2. Associativity
3. Existence of Identity Element
4. Existence of Inverse Element.

For 1.
It is clear that closure follows from the definition of the binary operation used with the quotient structure structure given that the product of two cosets $NaNb$ is another coset of the same parity Nab in the quotient structure, $\frac{G}{N} : a, b \in G$.

For 2.
We claim that for associativity to hold $(NaNb)Nc = Na(NbNc)$.
Consider LHS :

$$(NaNb)Nc = NabNc = Nabc$$

and RHS :

$$Na(NbNc) = NaNbc = Nabc.$$

Therefore $(NaNb)Nc = Na(NbNc)$ associativity holds.

For 3.
Consider that $Ne \in \frac{G}{N}$ and by definition of $e \in G$

$$NgNe = Nge = Ng \,\forall\, Ng \in \frac{G}{N} \qquad NeNg = Neg = Ng \,\forall\, Ng \in \frac{G}{N}.$$

Thus by definition of identity element Ne is the identity element in $\frac{G}{N}$.

For 4.
Consider that

$$\forall Ng \in \frac{G}{N} \,\exists\, Ng^{-1} \in \frac{G}{N} \;\because\; \forall\, g \in G \,\exists\, g^{-1} \in G$$

such that

$$NgNg^{-1} = Ngg^{-1} = Ne \qquad Ng^{-1}Ng = Ng^{-1}g = Ne$$

$$\therefore (Ng)^{-1} = Ng^{-1} \,\forall\, Ng \in \frac{G}{N}$$

Hence, given that all four axioms have been satisfied the quotient structure $\frac{G}{N}$ can be understood to be a group, the quotient group. □

It stands to reason, by Lagrange's Theorem, that Quotient Groups being made completely of cosets have order equal to the index of the normal subgroup in the arching group such that

$$\circ\left(\frac{G}{N}\right) = \frac{\circ(G)}{\circ N} = i_G(N).$$

An Example from Herstein Let G be the group of integers under addition and let the subgroup N be the set of all multiples of 3.
This subgroup is evidently normal given that for any integer $g \in G\, g^{-1}Ng = N$. For illustration consider

$$g^{-1}Ng = N \sim (-1) + 3 + 1 = 3.$$

We claim that this normal subgroup gives the subgroups $N, N+1, N+2$ exclusively.

Consider $N + a \ : \ a = 3b + c$ where $b \in G$ and c is the set of all possible remainders of a when dividing by 3. This gives 3 cases;

$$N + 3b \qquad N + 3b + 1 \qquad N + 3b + 2$$

Given that $3b$ is a multiple of 3 it is represented by N

$$N \qquad N + 1 \qquad N + 2$$

As a result the quotient group can be listed as

$$\frac{G}{N} = \{N, N + 1, N + 2\}.$$

Coset products in this case would follow the form

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

The Power of Quotient and Normal Subgroups Starting with some group G normal subgroups allow us to split a group into smaller groupings (cosets) whose elements are of the same type and quotient groups then allow us to abstract away, zoom out and think of the group in terms of the cosets and no longer in terms of elements. This is the scope of these two mathematical tools.

2.4.2 Homomorphisms

The coset product binary operation seems to preserve the structure of parity. As alluded to at the end of **Section 2.3.**, on Normal subgroups, a generalisation of this would allow us to create groups which share some property between their cosets, predicated by the normal subgroup and so by the kernel as it is known within this abstraction. This abstraction is called a homomorphism and is a mapping which preserves structure between algebraic structures, be they groups, rings or vector spaces.

Definition 2.4.2. A mapping ϕ from a group $(G, \cdot)$ into a group $(\bar{G}, *)$ is said to be a **homomorphism** if for all $a, b \in G, \phi(a \cdot b) = \phi(a) * \phi(b)$.

This means that to prove the existence of a homomorphism one must show well definition of some relation and that this relation satisfies the structural preservation; $\phi(ab) = \phi(a)\phi(b)$.

It is good to note that structure is being preserved, whilst not necessarily (ie: can be within the same group), across different groups and products and this is really the ubiquity and depth of the homomorphism.

Introductory Example : The Trivial Homomorphism

$$\phi(x) = e \, \forall \, x \in G$$

This mapping is inherently well-defined given that each element in the domain is uniquely mapped to the identity in the codomain.
Now the structure is also preserved by closure in the group given that all group elements are mapped to the identity in the codomain;

$$\phi(ab) = e \text{ and } \phi(a)\phi(b) = ee = e.$$

It stands to reason that such a mapping can be applied to any element in the codomain not just the identity, giving

$$\phi(x) = x \, \forall \, x \in G.$$

Example : Exponents of Real Numbers Let $G = (\mathbb{R}, +)$ and $\bar{G} = (\mathbb{R} \setminus 0, \times)$. Define

$$\phi(G) \to \phi(\bar{G}) : \phi(a) = 2^a.$$

Well-definition is inherent from the definition of exponents in $\mathbb{R} \to \mathbb{R} \setminus 0$.
The preservation of structure must be examined more closely. Consider that by laws of exponents

$$2^{a+b} = 2^a 2^b$$

which exposes readily

$$\phi(ab) = \phi(a)\phi(b)$$

such that $ab = a + b \in G = (\mathbb{R}, +)$ and $2^a \times 2^b \in \bar{G} = (\mathbb{R} \setminus 0, \times)$. Thus we indeed have a homomorphism.
This is an injective homomorphism as but it is not surjective given that only $\mathbb{R}^+ \setminus 0$ is made use of as the image by nature of 2^a.

Example : Matrices and Determinants Let $G = \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \forall\, a,b,c,d \in \mathbb{R}, \cdot \right)$ and $\bar{G} = (\mathbb{R} \setminus 0, \times)$. Define

$$f : G \to \bar{G} : \phi \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

First we note that f is nothing more than a linear combination of the $a, b, c, d \in \mathbb{R}$ and as such each matrix in G must be uniquely mapped to a determinant in $\bar{G}$. Thus f is a well-defined map.
Moving on we must show structural preservation; consider $A, B \in G$ with $a, b, c, d, e, f, g, h \in \mathbb{R}$:

$$\begin{aligned} A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad & B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}. \\ \implies \phi(A) = ad - bc \qquad & \phi(B) = eh - fg. \end{aligned}$$

Which gives

$$\begin{aligned} \phi(A)\phi(B) &= (ad - bc)(eh - fg) \\ \phi(A)\phi(B) &= adeh - adfg - bceh + bcfg \\ \phi(A)\phi(B) &= ad(eh - fg) + bc(fg - eh). \end{aligned}$$

Now consider AB;

$$\begin{aligned} AB &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} \\ AB &= \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix} \\ \implies \phi(AB) &= (ae + bg)(cf + dh) - (af + bh)(ce + dg) \\ \phi(AB) &= \cancel{aecf} + aedh + bgcf + \cancel{bgdh} - \cancel{afce} - afdg - bhce - \cancel{bhdg} \\ \phi(AB) &= ad(eh - fg) + bc(fg - eh) \\ \therefore\ \phi(AB) &= \phi(A)\phi(B). \end{aligned}$$

f is evidently a homomorphism.
The structural preservation proof could have been skipped by assuming $\det(AB) = \det(A)\det(B)$ by this was found to be some what not complete and given that it was specified for 2×2 matrices it begged for a worked out solution.
The Kernel of this Homomorphism is the Special Linear Group denoted $SL(n, \mathbb{R})$.

Example : The Dihedral Group and its normal subgroup Let $G = D_3 \simeq S_3$ and $\bar{G} = \{e, b\}$. Define

$$f : G \to \bar{G} : f(b^i a^i) = \phi^i.$$

First we note that $\bar{G} = \{e, b\} \leq G$ as $\bar{G}$ is evidently the reflection generating subgroup in D_3. Thus a mapping $f : G \to \bar{G}$ must be well-defined by this cardinality. Now for

the preservation of structure consider that any element in D_3 has the form $b^i a^j : b^2 = e$ and $a^3 = e$ and being a group features closure those if we consider two elements say $b^i a^j$ and $b^w a^z$ we get the product $b^{i+w} a^{j+z}$. With this in mind consider

$$f(b^i a^j) f(b^w a^j) = b^i b^w = b^{i+w}$$

and also

$$f(b^{i+w} a^{j+z}) = b^{i+w}$$

showing that $f(st) = f(s)f(t) : s, t \in D_3$.

Now from the earlier section on visualisation of normal subgroups it should be easy to appreciate that the only normal subgroup in D_3 is C_3 and from what was earlier covered on normal subgroups exposing cosets and giving the quotient structure from which homomorphisms are birthed we are able to see that

$$N = C_3 = \{e, a, a^2\} \text{ and } \frac{G}{N} = \{e, b\} = \bar{G}.$$

And so N vanishes from the structure within the process (Kernel) ! This is a very powerful result which we will review in the coming section on Noether's first isomorphism theorem.

The Canonical or Natural Homomorphism In **Section 2.4.1.** and in the introduction to **Section 2.4.2.** it was strongly emphasised the coset product relation was the starting point for homomorphisms. Let's formalise this idea and prove it.

Theorem 2.4.3. *Suppose G is a group, N is a normal subgroup of G; define the mapping $\phi : G \to \frac{G}{N} : \phi(x) = Nx \, \forall x \in G$. Then ϕ is a homomorphism of G onto $\frac{G}{N}$.*

Well-definition of the coset product was shown in **Lemma 2.4.1.**
The structure is preserved by the fact that parity is preserved within cosets products such that

$$\phi(ab) = Nab = NaNb = \phi(a)\phi(b).$$

This mapping is evidently homomorphic but we also claim that it is a surjective homomorphism. Consider an element in the codomain $X \in \frac{G}{N} : X = Ny$, $y \in G$ but given that this is a coset it implies that it is in the image of the homomorphism

$$X = \phi(y).$$

This implies that the image coincides with the whole domain. Other than this the mapping is also injective such that an element in the factor group is uniquely mapped to only one element in the domain.
This will prove to be a very particular type of homomorphism which will be discussed in the section on isomorphisms.

It is relevant to note that the factor group is $\bar{G}$ and that the normal subgroup has vanished from the group leaving only the factor group in the image. Again we are alluding to two new structures, one of which will be introduced hereunder. The kernel.

Kernels and their structure within Homorphisms Let's recall the example from the dihedral group for illustration. We were mapping from

$$f : G \to \bar{G}$$

$$f : \{e, a, a^2, b, ba, ba^2\} \to \{e, b\}.$$

Now the kernel is the name given to a part of the original structure which is lost within a homomorphic map.
Clearly C_3 is being lost on going from the domain to the image of the function as $\{e, b\}$ lacks products involving the cyclic elements and the elements itself. From our intuition gained through the Canonical homomorphism we claim that the kernel is the normal subgroup such that the image is the quotient group. We have just laid the foundations for Noether's First Isomorphism Theorem.

Now how is C_3 lost or how does it vanish? By mapping the elements of C_3 in D_3 to e in $\bar{G}$, since C_3 is the Kernel. This leads to the definition of the Kernel as a set.

Definition 2.4.3. If ϕ is a homomorphism of G into $\bar{G}$, the kernel of ϕ, $\ker(\phi)$ is defined by

$$\ker(\phi) = \{x \in G | \phi(x) = \bar{e}\}$$

such that $\bar{e}$ is the identity element in $\bar{G}$.

Lemma 2.4.4. *If ϕ is a homomorphism of G into $\bar{G}$, then*

1. *$\phi(e) = \bar{e}$, the unit element of $\bar{G}$ $\implies$ the kernel is non-empty.*
2. *$\phi(x^{-1}) = \phi(x)^{-1} \, \forall \, x \in G$*

Proof. For *1*.

$$\phi(x)\bar{e} = \phi(x) = \phi(xe)$$

Now since ϕ is a homomorphism

$$\begin{gathered}\phi(x)\bar{e} = \phi(x) = \phi(xe) = \phi(x)\phi(e) \\ \implies \phi(x)\bar{e} = \phi(x)\phi(e).\end{gathered}$$

By the cancellation property in $\bar{G}$

$$\phi(e) = \bar{e}.$$

For *2*.

Consider that by *1*. $\bar{e} = \phi(e)$.

By def'n of inverse in G

$$\phi(e) = \phi(xx^{-1}) \qquad\qquad \phi(e) = \phi(x^{-1}x).$$

Given that ϕ is a homomorphism

$$\phi(e) = \phi(xx^{-1}) = \phi(x)\phi(x^{-1}) \qquad \phi(e) = \phi(x^{-1}x) = \phi(x^{-1})\phi(x).$$

$$\implies \phi(e) = \phi(x)\phi(x^{-1}) = \phi(x^{-1})\phi(x)$$

$$\therefore \phi(x^{-1}) = \phi(x)^{-1} \text{ by definition of inverse.}$$

□

The Kernel is the Normal Subgroup It has been alluded to many times within this section that it is the normal subgroup in G which vanishes in $\bar{G}$. Nowhere was this more clear than within the canonical isomorphism wherein $\bar{G} = \frac{G}{N}$. Let us now formalise this notion, showing that it is so for any group homomorphism.

Theorem 2.4.5. *If ϕ is a homomorphism of G into $\bar{G}$ with kernel K, then K is a normal subgroup of G.*

Proof. The proof must show two things. First we show that $\ker(\phi)$ is indeed a subgroup in G and then show that it is normal.
Starting with the subgroup property we show closure and the existence of an inverse element for all subgroup elements.

For Closure;

$$\text{Let } a, b \in \ker(\phi)$$

Making use of the definition of $\ker(\phi)$

$$\implies \phi(a) = \bar{e} \text{ and } \phi(b) = \bar{e}.$$

$$\text{Consider } \phi(ab) = \phi(a)\phi(b) \text{ since } \phi \text{ is a homomorphism.}$$

$$\implies \phi(ab) = \bar{e}.$$

$$\therefore ab \in \ker(\phi) \text{ showing closure, as required.}$$

For Inverse Law by **Lemma 2.4.4. ii)**;

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

Thus $x^{-1} \in \ker(\phi)$ showing that it is a subgroup.

To show that this subgroup is normal we use the definition of normality

$$N \trianglelefteq G \iff g^{-1}Ng \subseteq N.$$

Meaning we mush show

$$g^{-1}\ker(\phi)g \subseteq \ker(\phi).$$

$$\begin{gathered}\text{Let } x \in g^{-1}\ker(\phi)g.\\ \implies x = g^{-1}kg \,:\, \exists\, k \in \ker(\phi)\\ \phi(x) = \phi(g^{-1}kg)\end{gathered}$$

Given that ϕ is a homomorphism

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

By definition of $\ker(\phi)$)

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

By definition of $\bar{e} \in \bar{G}$

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

By definition of inverse

$$\begin{gathered}\phi(x) = \bar{e}.\\ \implies x \in g^{-1}\ker(\phi)g \implies x \in \ker(\phi)\\ \implies g^{-1}\ker(\phi)g \subseteq \ker(\phi)\\ \therefore \ker(\phi) \trianglelefteq G\end{gathered}$$

□

What happens to elements in G that are not in the kernel ?

In the words of Prof.Sciriha all the constituents of a village are represented fully by their village.

Theorem 2.4.6. *If ϕ is a homomorphism of G onto $\bar{G}$ with kernel K, then the set of all inverse images of $\bar{g} \in \bar{G}$ under ϕ in G is given by Kx, where x is any particular inverse image of $\bar{g}$ in G.*

The domain group contains a normal subgroup which acts as the kernel, all of whose elements are mapped to $\bar{e}$ in $\bar{G}$ but given that there are other elements in the domain; which elements in the codomain map to them ?

Proof. Assume there is some $\bar{g}$ whose inverse map is x

$$\phi(x) = \bar{g} \neq \bar{e}.$$

Are there other inverse maps that can be formed from $\bar{e}$? Consider

$$y = kx \,:\, k \in K \implies \phi(y) = \phi(kx) = \phi(k)\phi(x) = \bar{e}x = x \,\forall k \in K$$

Therefore $\bar{g}$ has inverse maps throughout all of elements in the coset Kx including the initial ex. So structurally we have shown that cosets work but are they the only structure in G which offers inverse maps for $\bar{g}$. We thus have a uniqueness sub-proof on our hands.
Assume that there is some z which is another inverse map for $\bar{g}$

$$\begin{gathered}\text{Let } \phi(z) = \bar{g} = \phi(x).\\ \phi(z) = \phi(x)\\ \phi(z)\phi(x)^{-1} = \bar{e}\end{gathered}$$

Since ϕ is a homomorphism, then by **Lemma 2.4.4.** ii)

$$\phi(z)\phi(x^{-1}) = \bar{e}.$$

By definition of homomorphism

$$\begin{gathered}\phi(zx^{-1}) = \bar{e}.\\ \implies zx^{-1} = k \in K\\ zx^{-1}x = kx\\ ze = kx\end{gathered}$$

Thus by definition of identity

$$\begin{gathered}z = kx \in Kx\\ \therefore\ z \in Kx\end{gathered}$$

The coset Kx exhausts all the inverse map possibilities of said $\bar{g}$. □

With this theory in mind a special case comes to light. If $\ker(\phi) = e$ then only $e \to \bar{e}$ and any $\bar{g}$ will have an inverse image $ex = x$ resulting in an injective mapping. This gives the following definition;

Definition 2.4.4. A homomorphism ϕ from G into $\bar{G}$ is said to be an isomorphism if ϕ is one-to-one, an injective homomorphism.
Definition 2.4.5. Two groups G and G^* are said to be isomorphic if they are related by a bijective homomorphism. This is denoted $G \simeq G^*$.

It's easy to appreciate that

1. $G \simeq G$

2. $G \simeq G^* \implies G^* \simeq G$

3. $G \simeq G^* \wedge G^* \simeq G^{**} \implies G \simeq G^{**}$

Prof.Sciriha would often remark that isomorphic groups are like *"the same thing in diguise"* and this is very astute. The fact that a bijective function can relate the two structures infers that they are just the same structure relabelled in some different way. We encountered isomorphisms before in these notes when discussing cyclic groups and residue groups.

Corollary 2.4.1. *A homomorphism ϕ of G into $\bar{G}$ with kernel K_ϕ is an isomorphism if and only if $K_\phi = e$.*

This corollary exposes a useful method for proving the existence of isomorphisms by first showing that a homomorphism exists and then showing that its kernel consists solely of e. This method will be flushed out in the theorem to follow in **Section 2.4.3.**

Example - Isomorphic Cyclic Groups Two cyclic groups of the same order are isomorphic.

Proof. To show that there is an isomorphism we show well-definition, homomorphism, surjectivity and injectivity.

Let $G = \langle g \rangle$ and $H = \langle h \rangle$ be cyclic groups of the same order. We define a function from $\theta : G \to H$

$$\theta : G \to H : \theta(g^m) = h^m.$$

For well-defintition;
The two groups can either have infinite or finite order, say n. If they both have infinite order then all the powers of g are distinct and θ is well-defined. Now consider the finite case, we wish to show that an element in the domain is mapped to one element in the codomain, ie: it does not have an ambiguous value.

$$\begin{gathered}\text{Suppose } g^k = g^l.\\ g^k g^{-l} = e\\ g^{k-l} = e\\ \implies n|k-l\end{gathered}$$

Given that H is of order n also

$$\begin{gathered}\implies h^{k-l} = e.\\ \therefore h^k = h^l\end{gathered}$$

One element in the domain has an unambiguous value in the codomain.

For bijection;
Consider that given that the two groups are of the same order then the function must be

surjective. Injectivity is inherent of the definition of the function such that an element in G with integer order, a, is mapped to an element in H with integer order, a, also.

For structural preservation;

$$\theta(g^k g^l) = \phi(g^{k+l}) = h^{k+l} = h^k h^l = \theta(g^k)\theta(g^l).$$

□

Theorem 2.4.7. *The Image of G, $\phi(G)$, is a subgroup of $\bar{G}$ for some homomorphism $\phi : G \to \bar{G}$.*

Proof. This is a subgroup proof and as such we are to show closure and the existence of an inverse for each element.

For Closure;
Let $a, b \in \phi(G)$

$$\implies \exists\ g_1, g_2 \in G\ :\ \phi(g_1) = a \text{ and } \phi(g_2) = b.$$

By closure in G, $g_1 g_2 \in G$ and by nature of homorphisms

$$\implies \phi(g_1 g_2) = \phi(g_1)\phi(g_2) = ab$$

$$\therefore\ ab \in \phi(G).$$

Showing closure, as required.

For Inverse;
Consider the fact that an inverse exists in G and recall **Lemma 2.4.4. ii)**

$$\phi(g_1)\phi(g_1^{-1}) = \phi(g_1)\phi(g_1)^{-1} = \phi(g_1)^{-1}\phi(g_1) = \bar{e}.$$

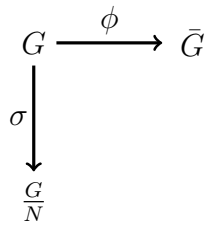
Therefore, $\phi(g)^{-1} \in \phi(G), \forall g \in G$ is the inverse element.
The Image of a homomorphism has shown to be a subgroup of the homomorphism □

2.4.3 Noether's First Isomorphism Theorem

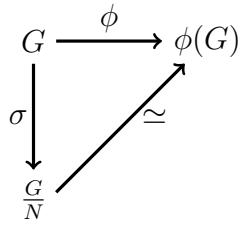
A good place to start is what we currently know, so let's show what we know for some groups G and $\bar{G}$ assuming that they are related by some homomorphism ψ. This would give;

$$G \xrightarrow{\ \psi\ } \bar{G}$$

Now we also investigated the fact that any Kernel is a normal subgroup and within the section on Quotient Groups as well as the Canonical Homomorphism we showed that we can turn a group into a group of its cosets by such a relation. Let's label this σ.



It seems only natural to wish to relate the factor group and the $\bar{G}$ in some way to complete the triangle. Consider that the ϕ maps to G to the image subgroup $\phi(G)$: $\phi(G) \leq \bar{G}$ thus we require a *surjective* homomorphism so that $\phi(G) = \bar{G}$. For such a case we consider that within both homomorphisms the Kernel is the same, the normal subgroup and $\bar{G}$ and $\frac{G}{N}$ appear to be *the same thing in disguise.*



This is our claim in the first isomorphism theorem that the image of a homomorphism is isomorphic to the quotient group and it is predicated on two theorems that we have proven;

1. $\phi(G)$ is a subgroup of H
2. $\ker(\phi)$ is a normal subgroup of G.

Theorem 2.4.8. *Let ϕ be a homomorphism of G to $\bar{G}$ with kernel K. Then $\frac{G}{K} \simeq \phi(G)$.*

Proof. We wish to show that $\frac{G}{K} \simeq \phi(G)$ so we must show that a bijective homomorphism can be formed between the two groups. Let's define a mapping

$$\psi : \frac{G}{K} \to \phi(G) \text{ where } x \in \frac{G}{K}, x = Kg \text{ then } \phi(x) = \phi(g).$$

To prove that such a mapping is isomorphic we must show, well-definition, structural preservation and bijectivity.

For well-definition;
Consider that an $X \in \frac{G}{K}$ may be represented by multiple cosets. We show that said representations have the same output in the image giving an unambiguous definition of the value for X.

$$\text{Let } X \in \frac{G}{K}$$
$$\implies X = Kg_1 = Kg_2 \ \exists \ g_1, g_2 \in G$$
$$\implies eg_1 = kg_2 \ \exists \ k \in K$$
$$\text{Consider } \psi(Kg_1) = \phi(eg_1) \text{ by defintion of } \psi.$$
$$\phi(eg_1) = \phi(kg_2)$$

Since ϕ is a homomorphism

$$\phi(e)\phi(g_1) = \phi(k)\phi(g_2).$$

Given that e and k are in K

$$\begin{gathered}\bar{e}\phi(g_1) = \bar{e}\phi(g_2).\\ \implies \phi(g_1) = \phi(g_2)\\ \therefore \psi(Kg_1) = \psi(Kg_2)\end{gathered}$$

Thus, ψ is well-defined as required given that an element in $\frac{G}{K}$ has one value in $\phi(G)$.

Moving on we show that the function is homomorphic;
Let $Kg_1, Kg_2 \in \frac{G}{N}$
By **Theorem 2.3.4.** the product of two right cosets is a right coset $Kg_1Kg_2 = Kg_1g_2$.
Now by definition of ψ and the fact that ϕ is a homomorphism

$$\psi(Kg_1)\psi(Kg_2) = \phi(g_1)\phi(g_2) = \phi(g_1g_2)$$

and from above

$$\psi(Kg_1Kg_2) = \psi(Kg_1g_2) = \phi(g_1g_2).$$

Therefore

$$\psi(Kg_1Kg_2) = \psi(Kg_1)\psi(Kg_2),$$

ψ is a homomorphism also.

For surjectivity;
Consider an element, z in the range $\phi(G)$;

$$\implies z = \phi(g) \ \exists \ g \in G.$$

By def'n of ψ

$$\psi(Kg) = \phi(g) = z$$

Thus showing that for any $\phi(g)$ there is a $Kg \in \frac{G}{K}$ and so the whole range is the codomain.

For injectivity;
We consider two elements in the codomain and show that if they are equal then they must map to the same element in the domain if the function is injective.
Let $\psi(Kg_1) = \psi(Kg_2)$ and by definition of ψ

$$\begin{gathered}\phi(g_1) = \phi(g_2).\\ \phi(g_1)\phi(g_2)^{-1} = \bar{e}\\ \phi(g_1)\phi(g_2^{-1}) = \bar{e}\end{gathered}$$

Since ϕ is a homomorphism

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

By def'n of K

$$g_1 g_2^{-1} \in K.$$
$$g_1 g_2^{-1} = k.$$
$$g_1 \in K g_2$$
$$k g_1 \in K g_2$$
$$K g_1 \subseteq K g_2$$

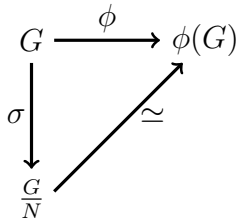
Given that cosets are either identical or disjoint

$$K g_1 = K g_2.$$

Thus two identical elements in the codomain map to the same element in the domain showing that the function is injective also and thus bijective.
Alternatively, injectivity can be proven showing that the kernel of the function from $\frac{G}{K} \to \phi(G)$ is the identity element only. Elements within kernels are mapped to $\bar{e}$ in the image resulting in a many-to-one function thus showing that the kernel only contains an identity element is sufficient for injectivity.

Given that we have a bijective homomorphism we have also proven that G is isomorphic to $\phi(G)$. □

The relational diagram has now been established relating a group, its image and its quotient group.



2.4.4 Cayley's Theorem

Groups are a type of algebraic structure and throughout the unit, the uninitiated may have thought that groups that are labelled differently or act in different contexts, say cyclic groups and residue groups are fundamentally different. But having now introduced the notion of isomorphisms we begin to appreciate a strange phenomenon linking all groups together. This is what Arthur Cayley noted and what this theorem is about. Cayley noted that the only thing which distinguished groups from one another other

than order and structure was their, context and labelling. For example, the aforementioned residue and cyclic groups show up in completely different settings. Within the first we're dealing with modular multiplication and coprimes, within the second we're rotating shapes but when you observe their structure, perhaps through a Cayley Table...

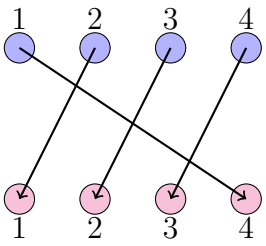
C_4	e	a	a^2	a^3
e	e	a	a^2	a^3
a	a	a^2	a^3	e
a^2	a^2	a^3	e	a
a^3	a^3	e	a	a^2

$x \mod 5$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

the groups begin to look similar. They both have a single generator and four elements so they are structurally identical and we can show this through an isomorphism but Cayley took this a step further. He noticed that by treating groups as permutation one can expose the fact that they are all doing the same thing, shuffling things about.
The group group concerned with shuffling is the Symmetric group ! Cayley claimed that any group is isomorphic to some subgroup of the Symmetric group.

Theorem 2.4.9. *Every group is isomorphic to a subgroup of S_n for some appropriate n.*

The proof will follow by using the tool Cayley himself used, the Cayley Table, to map each row corresponding to some $g \in G$ to a permutation.



So let's consider a practical example. Again we think about C_4, labelling each of its elements such that $e = 1, a = 2, a^2 = 3$ and $a^3 = 4$. Now we can start thinking about products as permutations; multiplying by e of course leaves everything the same so that is the starting $(1)(2)(3)(4)$ and now considering a we see that on multiplying by a the following shuffle occurs which can be represented as (4321) as a cycle.
Applying this idea to each row we get;

C_4	e	a	a^2	a^3	Cycle
e	e	a	a^2	a^3	$(1)(2)(3)(4)$
a	a	a^2	a^3	e	(4321)
a^2	a^2	a^3	e	a	$(13)(24)$
a^3	a^3	e	a	a^2	(1234)

This the central idea of Cayley's theorem. Let's formalise what we just did here ie: map group elements to permutations ex:

$$e \to (1)(2)(3)(4)$$

into a proper mapping and show that this is an isomorphic map.

Proof. Let G be a group and let $P_x : G \to G$ be defined by $P_x : g \to xg \,\forall\, x \in G$ where P_x is a permutation.

Note: $P_x : g \to xg$ makes sense as a permutation if you think about say *one goes to four* where *one* is representing e going to *four* by multiplying by *four* which is x in this case, a^3. Another exmaple, for *two goes to one* we multiply *two* by any number larger than *two* $\{3, 4\} = \{a^2, a^3\}$ which turns it to *one*, e.

Permutations are inherently bijective but we note that well-definition is inherited from the group operation, injectivity from $xg_1 = xg_2 \implies g_1 = g_2$ and surjectivity from the fact that the domain and the codomain are the same set.

We define a mapping ψ which relates a group element $g \in G$ to a row in its Cayley table then exposing it to be a permutation of group elements[1].

$$\psi : G \to \{P_G\}$$

defined by

$$P_g(h) = gh \,\forall\, h \in G.$$

We proceed to show that this mapping is well-defined, homomorphic and bijective.

Well definition of ψ is inherited from the well-definition of the binary operation in G since $P_g(h) = gh$ and gh is a group product.

For structural preservation;
Let $g_1, g_2 \in G$ and by definition of ψ we have

$$\psi(g_1) = P_{g_1}(h) \qquad\qquad \psi(g_2) = P_{g_2}(h).$$

$$\implies \psi(g_1)\psi(g_2) = P_{g_1}P_{g_2}(h) = g_1g_2h$$

Now considering

$$\psi(g_1g_2) = P_{g_1g_2}(h) = g_1g_2h$$

then

$$\psi(g_1g_2) = \psi(g_1)\psi(g_2).$$

Thus, ψ is a homomorphism as required.

For injectivity we take two elements in the codomain such that they are equal and show that if we have an injection then they must map to the same element in the domain.

$$\text{Let} \psi(g_1) = \psi(g_2).$$

By definition of ψ

$$g_1h = g_2h.$$
$$g_1hh^{-1} = g_2hh^{-1}$$

[1] But then how will we get the cycle form ? One should appreciate that a bijection is apparent between the rows of Cayley tables and permutations as shown on page 61

By definition of inverse

$$g_1 e = g_2 e.$$

By definition of identity

$$g_1 = g_2.$$

Showing injectivity as required.

For surjectivity we show that the range is the whole codomain by showing that for every element in the range there is an element in the domain.
Let P_x be a permutation in the codomain. This implies that $x \in G$ and $\psi(x) = P_x \, \forall \, x \in G$. Showing surjectivity as required.

Therefore having shown that

$$\psi : G \to \{P_G\}$$

is well-defined, homomorphic, injective and surjective;

$$G \simeq \{P_G\} \subseteq S_n.$$

□

And so we have proved Cayley's Theorem showing that whilst groups are indeed an abstract algebraic structure that may be generalised by four axioms there is in truth only one group. This should not disparage us as it is the mathematical context and not the structure which makes a group not only interesting, but useful.

Chapter 3

Vector Spaces

3.1 Introduction and Formalisms

Up until this point one's experience may have been limited to playing around with 3-D Euclidean space in units like Analytical Geometry and Classical Mechanics. Now that Groups have taught us about what it means to be an abstract algebraic object we can apply the same approach to Vector Spaces seeing that they too are a general mathematical object with axioms and qualities which will allow us to extend their concepts infinitely many dimensions and different cases.

What's a Vector Space ? Vector Spaces are defined over *fields* which are a special class of *rings* thus before describing groups we must venture to outline, in brevity given the scope of these notes, the nature of rings and and fields.

Rings are built off of groups such that just like groups were sets with some binary operation, in rings we now have **sets with two binary operations**. In particular we have addition and multiplication where under addition an *associative ring*, which are those we are concerned with, forms an Abelian Group with addition and a normal group with multiplication.
To summarise in a more axiomatic and formal manner we can say
Definition 3.1.1. A nonempty set R is said to be an *associative ring* if in R there defined two operations, denoted by $+$ and $\cdot$ respectively : $\forall\, a, b, c \in R$:

1. $a + b \in R$
2. $a + b = b + a$
3. $(a + b) + c = a + (b + c)$
4. $\exists\ 0 \in R : a + 0 = a, \forall\, a \in R$
5. $\exists\ -a \in R : a + (-a) = 0, \forall\, a \in R$
6. $a \cdot b \in R$

7. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
8. $a \cdot (b + c) = a \cdot b + a \cdot c$ and $b \cdot a + c \cdot a$

It is evident from groups that axioms 1-5 show that $(R, +)$ is an Abelian Group and 6-7 show that $(R, \cdot)$ is a group. 8 are the distributive laws which shows how the two binary operations interact together.

Now Vector Spaces in particular make use of a particular class of rings known as **fields** which are **commutative division rings**.
This means that R does not only form an Abelian group with $(+)$ but also with $(\cdot)$. This is possible when one excludes 0 from R so that multiplicative inverses hold for all $a \in R$. Thus we have

$$F = \{(R, +) \text{ and } (R \setminus 0, \cdot)\}.$$

Having defined what a field is we're in good shape to define what vector spaces are.
Definition 3.1.2. Let F be a field. A vector space over F, is a set V with a binary operation $+$ and, for each $c \in F$, a unary operation of scalar multiplication by c, satisfying the following axioms:

1. **Closure Law :** $\forall\, v.w \in V,\ v + w \in V.$
2. **Associative Law :** $u + (v + w) = (u + v) + w \,\forall u, v, w \in V.$
3. **Identity Element :** $\exists\ 0 \in V : v + 0 = 0 + v = v \,\forall\, v \in V.$
4. **Inverse Law :** $\forall\, v \in V,\ \exists\ w \in W : v + w = w + v = 0 \,\forall\, v, w \in V.$
5. **Commutativity :** $v + w = w + v,\ \forall\, v, w \in V.$

These 5 axioms stem from the Abelian Group formed by $(V, +)$ as mentioned in the definition of rings.

1. **Closure Law :** $\forall\, c \in F$ and $v \in V$, $cv \in V$.
2. **Unital Law :** $\forall\, v \in V, 1 \in F,\ 1v = v$: 1 is the identity.

These 2 axiom stem from the structure of the of the unary product side of the commatitive division ring over which the field is defined.
The last 3 axioms make use of the whole structure ie: both the binary and unary operations.

1. **Distributivity M over A :** $\forall\, v, w \in V$ and $c \in F$, $c(v + w) = cv + cw$
2. **Distributivity A over M :** $\forall\, v \in V$ and $c, d \in F$, $(c + d)v = cv + dv$
3. **Associativity :** $\forall\, v \in V$ and $c, d \in F$, $(cd)v = c(dv)$

If we are particularly astute and take a page out of groups one can see that Axiom 1 looks a lot like a homomorphism meaning that c is a homomorphism from the vector space on to itself, an isomorphism ! This digression can be extended to endomorphism rings in vector spaces and modules but such a discussion is premature within the scope of these notes.

3.1.1 Examples

N-Tuples : - Coordinate Wise Let $V = F^n$, the set of all n-tuples[1] of elements of F.
Equality of elements will be defined

$$\forall\, a, b \in V,\ a = b \iff (a_1, a_2, a_3, \ldots, a_n) = (b_1, b_2, b_3, \ldots, b_n) \iff a_i = b_i\, \forall i \in [n].$$

Addition and scalar multiplication will be defined

$$(a_1, a_2, a_3, \ldots, a_n) + (b_1, b_2, b_3, \ldots, b_n) = (a_1 + b_1, a_2 + a_2, a_3 + a_3, \ldots, a_n + b_n)$$

and

$$c(a_1, a_2, a_3, \ldots, a_n) = (ca_1, ca_2, ca_3, \ldots, ca_n) \text{ for } c \in F.$$

It is easy to see that all 10 axioms will hold for this example, primarily making use of the inner entries and the fact that they are elements of F, a ring.
Such a representation of a vector space will prove to be very important for the study of finite dimensional vector spaces and is denoted by $F^{(n)}$. Do appreciate that this example reduced to 3 Dimensions will elucidate our approach to Euclidian Space in $\mathbb{R}^3$ such that a vector $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ is nothing more than an alternate representation of the n-tuple $(3, 2, 1)$.

N-Tuples : - Functions Pointwise Let Ω be any set, and let V be the set of all functions from Ω to the field F such that said functions make use of the field operations in the following ways

$$(f + g)(x) = f(x) + g(x)$$
$$(cf)(x) = cf(x).$$

Now for a finite $\Omega = \{x_1, \ldots, x_n\}$ we can represent a whole function by putting each of its points in an n-tuple giving

$$(f(x_1), \ldots, f(x_n)).$$

Which reduces this case to the earlier example showing an equivalence between functions and vectors in some sense.
For an infinite Ω say $\mathbb{R}$ things get a bit more complicated, but that's functional analysis.

3.1.2 Properties of Vector Spaces

Theorem 3.1.1.

1. ***Scalar*** $\times$ 0***-Vector :*** $\alpha 0 = 0$ *for* $\alpha \in F$, $0 \in V$.

[1] ordered sets with n entries

2. ***Vector*** $\times$ 0***-scalar :*** $0v = 0$ *for* $0 \in F$, $0, v \in V$.
3. ***Scalar*** > 0 ***:*** $(-\alpha)v = -(\alpha v)$ *for* $\alpha \in F$, $v \in V$.
4. ***Scalar*** $= 0$ ***:*** *If* $v \neq 0$, *then* $\alpha v = 0$ *implies that* $\alpha = 0$ *for* $\alpha, 0 \in F$, $v \in V$.
5. ***Vector*** $= 0$ ***:*** *If* $\alpha \neq 0$, *then* $\alpha v = 0$ *implies that* $v = 0$ *for* $\alpha \in F$, $v, 0 \in V$.

Proof. For *1.*;
Consider that by definition of 0-vector as identity

$$\alpha 0 = \alpha(0 + 0).$$

Now by distributivity of scalar multiplication over addition

$$\alpha 0 = \alpha 0 + \alpha 0 \implies \alpha 0 = \alpha 0 - \alpha 0.$$

And so by definition of the identity

$$\therefore \alpha 0 = 0.$$

For *2.*;
Consider that since $0 = 0 + 0$ in F

$$0v = (0 + 0)v.$$

Now by distributivity of addition over scalar multiplication

$$0v = 0v + 0v \implies 0v = 0v - 0v.$$

$$\therefore 0v = 0.$$

For *3.*
Consider that $\alpha v + (-\alpha)v = v(\alpha - \alpha)$ by reversal of distributivity. Now, since by *2.* and given that $-\alpha$ is the additive inverse of α

$$\alpha v + (-\alpha)v = 0v = 0.$$

Therefore by definition of additive inverse in $(V, +)$

$$-\alpha v = (-\alpha)v.$$

For *4.*;
Suppose for contradiction that $\alpha \neq 0$ also, this implies that $a^{-1} \in F$..
Now $\alpha v = 0$ but pre-multiplying by a^{-1} we have

$$a^{-1}av = a^{-1}0.$$

By definition of identity in $(F \setminus 0, \times)$ and *1.*

$$1v = 0$$

$$\therefore v = 0 \text{ ※}.$$

Therefore $\alpha = 0$ indeed.

For *5.*
Following the same procedure as within the proof of *4* we arrive to the same conclusion that $v = 0$ only this time it is what we required. □

3.2 Linear Dependence, Span and Basis

Vector Spaces as described, in the previous section, are sets of objects defined over some field, a commutative division ring. Now recall back to $\mathbb{R}^3$, it was clear that vectors are split into two distinct classes. Those which we use to build other vectors and thus exist on their own merit so to speak and are thus independent and those that are built by these vectors and are thus dependent.

3.2.1 Linear Dependence

This notion of building in vector spaces is formalised using linear combinations from linear algebra

Definition 3.2.1. If V is a vector space over F and if $v_1, \ldots, v_n \in V$ then any element of the form $a_1v_1 + a_2v_2 + \cdots + a_nv_n$, where the $a_i \in F$ is a linear combination over F of $v_1, \ldots, v_n$

Informally we can say that any vector which can be represented as a linear combination of some other vectors is linearly dependent. *This result will be proven momentarily.* And so cannot be one of vectors that build or generate the vector space.

As an example consider the following in $\mathbb{R}^3$ as we know it and some vector $\gamma(i+j)$ in an attempt to prove that it is dependent on $i+j$ and $i+2j$ with $\Omega = \{i+j+k\}$. Now from our relation between linear dependence and expression in terms of linear combination this means that $\exists\ \alpha, \beta \in \mathbb{Z}$ we have

$$\gamma(i+k) = \alpha(i+j) + \beta(i+2).$$

Now let's compare coefficients

$$\begin{aligned} i &: \gamma = \alpha + \beta \\ j &: 0 = \alpha + 2\beta \\ k &: \gamma = 0. \end{aligned}$$

And back substituting we get $\begin{cases} \gamma = 0 \\ \beta = 0 \\ \alpha = 0. \end{cases}$ This is suggestive of the formal definition for linear dependence of vectors as in this example the scalars are all zero and so $\gamma(i+j)$ cannot be represented a linear combination of $i+j$ and $i+2j$ and is thus linearly independent.

Definition 3.2.2. If V is a vector space and if $v_1, \ldots, v_n$ are in V, we say that they are **linearly dependent** over F if there exist elements $\lambda_1, \ldots, \lambda_n \in F$ with not all of them 0, such that

$$\lambda_1v_1 + \lambda_2v_2 + \cdots + \lambda_nv_n = 0.$$

Now this definition makes sense in a very pragmatic way because the dependence is evident if one were to leave one product on the LHS and take all the rest to RHS.

Turning this on its head we can define linear independence as follows
Definition 3.2.3. If V is a vector space and if $v_1, \ldots, v_n$ are in V, we say that they are **linearly independent** over F if there exist elements $\lambda_1, \ldots, \lambda_n \in F$ with ALL of them 0, such that

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0.$$

To compare these two definitions, within the dependent case; we need the scalars not all be zero so that when taking products to the opposing sides we have true equality given that such vectors can be decomposed to the same thing. Alternatively, the only way to get a 0 on RHS for such a case is if the vectors on LHS depend on each other and thus may give cancellation.
Within the independent case; we need all the scalars to have a value of 0 because when taking any product to the other side we would never have equality between LHS and RHS otherwise.

With these definitions in hand let us now show the equivalence of vectors being expressible as linear combinations and being linearly dependent.
Theorem 3.2.1. *Let $v_1, v_2, \ldots, v_k$ form a linearly dependent ordered set of non-zero vectors; then $\exists\, k : v_k$ is a linear combination of its predecessors*

The proof will proceed by first using the definition of linear dependence, mainly the fact that not all scalars will have a value of zero and that non-zero scalars have inverses which will allow for linear combination representations.

Proof. By definition of linear dependence we have

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0.$$

such that not all $\lambda = 0$.
Let k be the largest index of a non-zero scalar

$$\lambda_k \neq 0.$$

Given that F is commutative division ring this means that

$$\lambda_k^{-1} \in F.$$

Now by premise $\forall\, i > k$, $\lambda_i = 0$

$$\implies \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = 0$$

and since $\lambda_k^{-1} \in F$ then we can represent v_k as

$$v_k = -\lambda_k^{-1}(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{k-1} v_{k-1}) = (-\lambda_k^{-1}\lambda_1)v_1 + (-\lambda_k^{-1}\lambda_2)v_2 + \cdots + (-\lambda_k^{-1}\lambda_{k-1})v_{k-1}$$

which is evidently a linear combination of its predecessors. □

Remark 3.2.1. Any set of vectors containing the zero vector is linearly dependent.

3.2.2 Spanning Sets and Basis

Definition 3.2.4. If $V = \{v_1, v_2, \ldots, v_n\}$ is defined over a field F then the set of all linear combinations of these vectors is know as the **span** written $\langle v_1, \ldots, v_n \rangle_F$. The set of vectors which generates the set of all linear combinations is known as the **spanning set**.

Spanning sets are collections of vectors which are capable of generating all the vectors in the span of some vector space but consider that a spanning set can contain the exact amount of vectors needed to span the space or it may contain

Lemma 3.2.2. *If $S = \{v_1, v_2, \ldots, v_q\}$ is a linearly dependent spanning set for $V(F)$ then there exists some $v_k : S \setminus \{v_k\}$ is still a spanning set for $V(F)$.*

Proof. By the premise the vector space $V(F) = \langle v_1, v_2, \ldots, v_q \rangle$ which means that some vector in the span $v \in V$ can be represented

$$v = \sum_{i=1}^{q} \alpha_i v_i$$

by the definitions of span and spanning set.
Now also by the premise, the spanning set is linearly independent meaning that there is some v_k, in the spanning set, which can be represented as the sum of its predecessors in the spanning set

$$v_k = \sum_{j=1}^{k-1} \beta_j v_j \ : \ \beta \in F, v_j \in S.$$

This means that we can now reformulate v;

$$v = \sum_{i=1}^{k-1} \alpha_i v_i + \alpha_k v_k + \sum_{i=k+1}^{q} \alpha_i v_i$$

$$v = \sum_{i=1}^{k-1} \alpha_i v_i + \alpha_k \left(\sum_{j=1}^{k-1} \beta_j v_j \right) + \sum_{i=k+1}^{q} \alpha_i v_i.$$

Now given that the scope of the two summations is the same we may assimilate

$$v = \sum_{i=1}^{k-1} \left(\alpha_i + \alpha_k \beta_i \right) v_i + \sum_{i=k+1}^{q} \alpha_i v_i.$$

Therefore v is a linear combination of the vectors in $\{v_1, \ldots, v_{k-1}\} \cup \{v_{k+1}, \ldots, v_q\}$.

$$\therefore S \setminus \{v_k\} \text{ spans } V(F)$$

□

Basis and its alternate formulations At this point it has become clear that spanning sets can have some extra vectors which make then linearly dependent. And so there must be some minimal spanning set of linearly independent vectors which span the vector space. This notion called the basis of a vector space.
Definition 3.2.5. A list of vectors in some Vector Space, $V(F)$, are referred to as a **basis** if they are linearly dependent and span the vector space.
Remark 3.2.2. The **size** of the basis is called the **dimension** of the vector space.

Theorem 3.2.3. *The following criterions for a finite subset X of a vector space $V(F)$, to be basis are equivalent*

(a) X is a linearly independent spanning set.

(b) X is a minimal spanning set.

(c) X is a maximal linearly independent set.

Proof. We will show that (a) = (b) and (a) = (c) which would prove that all three statements are equivalent. It should be noted that an equivalence statement is a bi-conditional statement and will be proven as such.

For (a) = (b), first ($\implies$);
Let $X = \{x_1, x_2, \dots, x_n\}$ be a basis such that it generates $V(F)$ and

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = 0 : \forall\, \lambda \in \{\lambda_i\},\ \lambda = 0.$$

Suppose for contradiction that it is not minimal meaning there is some x_k which can be removed $\exists\, x_k : X \setminus \{x_k\}$ still spans $V(F)$. But this means that x_k can be expressed as a linear combination of $X \setminus \{x_k\}$ which would mean that X_k is linearly dependent ※.
For (a) = (b), secondly ($\impliedby$);
Let $X = \{x_1, x_2, \dots, x_n\}$ be a minimal spanning set of $V(F)$ and for contradiction suppose that is not linearly independent, and so linearly dependent.
This would mean that by **Lemma 3.2.2.** there is some x_k which can be removed such that $X \setminus x_k$ is still a spanning set ※.
Therefore a minimal spanning set is a basis.

For (a) = (c), first ($\implies$);
Let $X = \{x_1, x_2, \dots, x_n\}$ be a basis such that it generates $V(F)$ and

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = 0 : \forall\, \lambda \in \{\lambda_i\},\ \lambda = 0.$$

Suppose for contradiction that it is not maximal meaning there is some $x \in V$ which can be added to the spanning set such that $X \cup \{x\}$ is now the set. But by the premise X is a basis and generate all of V including X meaning that x is a linear combination of the vectors in X and that $X \cup \{x\}$ is linearly dependent ※.
For (a) = (c), secondly ($\implies$);
Let $X = \{x_1, x_2, \dots, x_n\}$ be a maximal, linearly independent set and suppose for contradiction that it is not a spanning set of $V(F)$.

This would infer that there is some $v \in V(F)$ that cannot be represented as a linear combination of the vectors in X and so can be added to X, giving $X \cup \{v\}$ which would still be linearly independent ※. Therefore a maximal linearly independent set is a basis.

Thus, the three definitions are equivalent. □

3.3 Steinitiz Replacement

Theorem 3.3.1. *Let $Y = \{y_1, y_2, \ldots, y_r\}$ be linearly independent vectors and let $X = \{x_1, x_2, \ldots, x_n\}$ be a linearly dependent spanning set. Then $r \leq n$.*

This theorem will show an algorithmic approach to extending a set of linearly independent vectors to a basis and then compare the size of the initial basis and the subset.
The proof will follow by affixing elements from Y with the basis X such that this results in a linearly dependent set from which we can detract some element from the basis elements given its contribution is now being carried out by the element from Y. In this way we riffle through Y swapping out elements of X until we end with a new basis and comparing how many xs we added and how many ys we lost we achieve the result required.

Proof. Considering a basis X and a linearly independent subset Y.
Let $Z_1 = \{y_r, x_1, x_2, \ldots, x_n\}$ be ordered, having affixed a vector from Y to the basis. This means that since $y_r \in Y \subseteq V$, Z_1 is linearly dependent. By **Lemma 3.2.2.** this infers that we can remove some element in Z_1, which can be expressed as linear combination, and still have a spanning set. We demarcate this element as x_k giving $U_1 = Z_1 \setminus \{x_k\}$, which is a basis.
Repeating this process, we add a vector from Y to the newly acquired basis Z_1, relabelling if necessary, giving

$$Z_2 = (Z_1 \setminus \{x_k\}) \cup \{y_{r-1}\} = \{y_{r-1}, y_r, x_1, x_2, \ldots, x_n\}.$$

Y_1 is evidently linearly dependent and again we will have some x_t which can be expressed as a linear combination of its predecessors which can be removed to yield another basis, $U_2 = Z_2 \setminus \{x_t\}$.
This process can be repeating exhausting all but the first term of Y having swapped $r - 1$ vectors, giving

$$U_{r-1} = \{y_2, \ldots, y_{r-1}, y_r, x_1, \ldots, x_{n-(r-1)}\}.$$

Now to finish this process we affix y_1 to U_{r-1} thus having shuffling all of the initial linearly independent set through the basis elements. Let's order this set giving

$$Z_n = \{y_1, y_2, \ldots, y_r, x_1, \ldots, x_{n-(r-1)}\}$$

which is linearly dependent given the addition. Now currently we have $n + 1$ elements and to achieve a basis we must remove one element from $X_{[n-(r-1)]}$ which can be represented as a linear combination of its predecessors. Now in $X_{[n-(r-1)]}$ we can have vectors

which are linearly independent from the vectors in Y but certainly it is non-empty and we have at least 1 element which we are to swap out to achieve n;

$$n-(r-1)\geq 1$$

$$n+1-r\geq 1$$

$$n\geq r \text{ as required.}$$

□

We have thus shown that a linearly independent subset in a vector space is either a basis or can be extended to a basis but in a broader sense. Let's formalise this in this context.

Basis Formation - Spanning Set perspective

Corollary 3.3.1. *A linearly independent set of vectors is either a spanning set or can be extended to a spanning set.*

Here we are saying that if some subset of a vector space is linearly independent then it either already forms a basis and is hence maximal or can be made maximal.
We take an algorithmic approach to finding what need be added to this subset by considering the composition of this subset with the known basis of the vector space. One now proceeds through this set union checking for linearly dependent vectors such that if they are part of the original subset they are put back but if they are in the basis they are removed if they are linearly dependent or added if they are linearly independent they are left in the union to extend the original subset to a basis.

Proof. Let $Y=\{y_1,y_2,\ldots,y_r\}$ be a linearly independent subset of a vector space, $V(F)$, and $B=\{x_1,x_2,\ldots,x_n\}$ be the basis of the vector space.
We construct the union of these two sets

$$Y\cup B=U=\{y_1,y_2,\ldots,y_r,x_1,x_2,\ldots,x_n\}.$$

For the trivial case this set is still linearly independent given that the union reduces to the basis given that we have a set of doubles so to speak, which given the properties of sets reduces to the basis.
For the non-trivial case; this set is evidently linearly dependent containing the basis and the subset of vectors. We now order this set and call it U_0 and move through it until we find some x_k which can be expressed in terms of the vectors from Y and removing this x_k will still give a spanning set by **Lemma 3.2.2.**. Label this set

$$U_1=Y\setminus\{x_k\}$$

and repeat the process until some U_s is obtained which is a linearly independent maximal set. The extended original subset which now acts as a basis. □

3.3.1 Further Basis and Dimension

Theorem 3.3.2. *Any two bases in the same vector space have the same number of elements.*

Proof. Consider $A = \{a_1, \ldots, a_{r_1}\}$ and $B = \{b_1, \ldots, b_{r_2}\}$. Now by Steinitz Replacement Theorem it is understood that $r_1, r_2 \leq n$, but A is a basis by the premise;

$$|B| \leq |A|.$$

But also B is a basis by the premise;

$$|A| \leq |B|.$$

Therefore it is evident that

$$|A| = |B|.$$

□

This infers that the dimension of a vector space is well-defined.
Theorem 3.3.3. *If V is an n-dimensional vector space then any linearly independent subset of V of size n is a basis.*

Proof. This seeming obviously true, an approach by contradiction shall be taken.
Let $Y \subseteq V : Y = \{y_1, \ldots, y_n\}$ and

$$\alpha_1 y_1 + \cdots + \alpha_n y_n = 0$$

such that all $\alpha_i = 0$ and assume it is not a basis.

This would infer that $\exists\, v \in V$ which cannot be expressed as a linear combination of $\{y_1, \ldots, y_n\}$ and that so $Y \cup \{v\}$ is linearly independent.
But this would mean that a linearly independent subset of V has size larger than the dimension of the vector space; $n+1 > n$ ※. Therefore the theorem holds by the Steinitz Replacement Thoerem. □

Theorem 3.3.4. *In an n-dimensional vector space, a spanning set of size n forms a basis.*

Again taking an approach by contradiction.
Let V be an n-dimensional vector space with a spanning set $Y = \{y_1, \ldots, y_n\}$ and assume it does not form a basis.

By the premise Y is linearly dependent meaning that there is some $y_k \in Y$ which can be removed and still yield a spanning set U_1 by **Theorem 3.2.2.**.

$$U_1 = Y \setminus \{y_k\}.$$

This process is repeated going through all of Y until it is linearly independent and thus a basis.
But given that we started with a spanning set of size n

$$|U_r| < r$$

meaning that the size of the basis is smaller than the dimension of the vector space ※.
Theorem 3.3.5. *B is a basis if and only if every vector in V is **uniquely** expressible in terms of the basis vectors.*

Proof. For ($\implies$);
Suppose for contradiction that B is a basis but there exists some $v \in V$ which can be represented as a linear combination of the vectors in B in different ways.

$$\begin{cases} v = \alpha_1 b_1 + \cdots + \alpha_n, b_n \\ v = \beta_1 b_1 + \cdots + \beta_n, b_n \end{cases}$$

Subtracting the two we get the following

$$(\alpha_1 - \beta_1)b_1 + \cdots + (\alpha_n - \beta_n)b_1 = 0.$$

Now given that B is a basis and thus a linearly independent spanning set, we have

$$\alpha_i - \beta_i = 0 \forall \alpha_i, \beta_i \in F$$

$$\implies \alpha_i = \beta_i \forall \alpha_i, \beta_i \in F.$$

Therefore this result gives

$$\alpha_1 b_1 + \cdots + \alpha_n, b_n = \beta_1 b_1 + \cdots + \beta_n, b_n.$$

And so v is uniquely represented ※.

For ($\impliedby$);
Approaching directly, let $v \in V$ be uniquely expressibly in terms of the vectors of B for all $v \in V$. This means that B spans V. We must thus show that B is linearly independent.

Consider some vector $y \in V : y = 0$, given that B spans V we have

$$\alpha_1 b_1, \ldots, \alpha_n b_n = 0.$$

The zero vector can be expanded by $0x = 0$ to give

$$\alpha_1 b_1, \ldots, \alpha_n b_n = 0 b_1 + \cdots + 0 b_n.$$

But by the premise of ($\impliedby$), y is uniquely represented in B which means

$$\alpha_i = 0 \, \forall \alpha_i \in F.$$

And so by definition □

3.4 The Dimension Theorem

3.5 Linear Transformations