Math 451: Abstract Algebra I

Wolmer V. Vasconcelos

Set 1

Fall 2009

Outline

- Pre-requisites: By permission only
- Textbook: See Syllabus
- webpage:www.math.rutgers.edu/(tilde)vasconce
- email : vasconce AT math.rutgers.edu
- Office hours [H228]: TF 2:4, or by arrangement

General Syllabus

- Composition Laws
- Groups, Subgroups, Homomorphisms
- Quotient Groups
- Groups of Symmetry
- Group Actions on Sets
- Representation Theory
- Basics of Rings
- The X-Topic

Outline

Syllabus

- Text: Michael Artin's *Algebra*, Prentice Hall, 1991, ISBN 0-13-004763-5. Note: Students may be able to obtain used copies online through addall.com or other websites.
- Prerequisites: By permission only. Typically requires an introductory abstract algebra course (very comfortable with complex numbers, for one), solid linear algebra (e.g. Math350) and appreciate the joys of proofs (Math300).
- Meeting times: MTh 3rd period (12:00–1:20), ARC-105, Busch Campus
- Final Exam: TBA

Topics

1. We will cover selections from the first 10 chapters of Artin's Algebra.
2. We will cover most or all of chapters 2 (basic group theory), 5 (groups and symmetry), 6 (more group theory including the Sylow theorems), 9 (group representation theory), and 10 (basics of rings).
3. Although we will be focusing on group theory, we may treat the chapters on linear algebra (1, 3, 4, and 7) as needed.
4. The term grade will be based on the results of the examinations, homework problems, and class participation. We will have 2 midterm (80 minute) exams and a final exam. Homework sets will be assigned approximately once per week. Weights: HW=100, H1=100, H2=100, F=200.

Outline

Composition Laws

The basic notion of the course is that of a composition law on a set $\mathbb{X}$.

A composition on a set $\mathbb{X}$ is a function assigning to ordered pairs of elements of $\mathbb{X}$ an element of $\mathbb{X}$,

$$(a, b) \mapsto \mathbf{f}(a, b).$$

That is a function of two variables on $\mathbb{X}$ with values in $\mathbb{X}$.

The interesting composition laws gives rise to the algebraic structures denoted by the set $\{\mathbb{X}, \mathbf{f}(\cdot, \cdot)\}$.

It is nicely represented in a composition table

$\mathbf{f}$	$*$	b	$*$
$*$	$*$	$*$	$*$
a	$*$	$\mathbf{f}(a, b)$	$*$
$*$	$*$	$*$	$*$

We represent it also as

$$\mathbb{X} \times \mathbb{X} \xrightarrow{\mathbf{f}} \mathbb{X}$$

The functional notation $\mathbf{f}(a, b)$ is often replaced by more suggestive or familiar notations:

$$\mathbf{f}(a, b) \to a \cdot b, a + b, a \circ b, a \times b, ab$$

Of course, attention must be paid to the fact that $\mathbf{f}(a, b)$ may be different from $\mathbf{f}(b, a)$.

Example

Let **A** be a nonempty set and let $\mathbb{X}$ be the set of all maps $\alpha : \mathbf{A} \to \mathbf{A}$,

$$\mathbb{X} = \mathrm{Maps}(\mathbf{A}, \mathbf{A}).$$

Composition gets its name from

$$(\alpha, \beta) \in \mathbb{X} \times \mathbb{X} \to \alpha \circ \beta \in \mathbb{X},$$

$$(\alpha \circ \beta)(a) = \alpha(\beta(a)), \quad a \in \mathbf{A}$$

Exxample

A set **A** with two elements, say $\mathbf{A} = \{0, 1\}$, has $\mathbb{X} = \{i,\ \tau, \alpha, \beta\}$, with i the identity map, τ the transposition map, that is $\tau(0) = 1$, $\tau(1) = 0$. α and β are the constant maps: $\alpha(1) = \alpha(0) = 0$ and $\beta(1) = \beta(0) = 1$.
The composition table is

$\circ$	i	τ	α	β
i	i	τ	α	β
τ	τ	i	β	α
α	α	α	α	α
β	β	β	β	β

Properties of Compositions

Let **f** be a composition on the set $\mathbb{X}$.

- **f** is associative if $\mathbf{f}(\mathbf{f}(a,b),c) = \mathbf{f}(a,\mathbf{f}(b,c))$ for all $a,b,c \in \mathbb{X}$. In the more standard notation:

$$a(bc) = (ab)c, \quad \forall a,b,c \in \mathbb{X}$$

- **f** is commutative if $\mathbf{f}(a,b) = \mathbf{f}(b,a)$ for all $a,b \in \mathbb{X}$,

$$ab = ba, \quad \forall a,b \in \mathbb{X}$$

- Note that the composition table of a commutative composition law is symmetric.
- For the composition law (natural?) defined above on $\mathbb{X} = \mathrm{Maps}(\mathbf{A},\mathbf{A})$, we have

$$\alpha \circ (\beta \circ \gamma) = (\alpha \circ \beta) \circ \gamma$$

since, as maps, they both equal $\alpha(\beta(\gamma(a)))$, $\forall a \in \mathbf{A}$.

Properties of Compositions

- An identity of a composition $\mathbf{f}$ is an element $e \in \mathbb{X}$ such that

$$\mathbf{f}(a, e) = \mathbf{f}(e, a) = a, \quad \forall a \in \mathbb{X}$$

  If an identity for $\mathbf{f}$ exists, it is unique: $e' = ee' = e'e = e$.
- For a compostion $\mathbf{f}$ with identity e, an inverse of an element $a \in \mathbb{X}$ is an element $b \in \mathbb{X}$ such that

$$\mathbf{f}(a, b) = \mathbf{f}(b, a) = e$$

  If $\mathbf{f}$ is associative, inverses, when they exist, are unique.
- If $\mathbf{f}$ is associative with identity, c is the inverse of a and d is the inverse of b, then $\mathbf{f}(d, c) = dc$ is the inverse of $\mathbf{f}(a, b) = ab$.

Properties of Compositions

- If **f** is associative, we may define powers: $a^1 = a$, and for $n > 1$ $a^n = a^{n-1}a$. Moreover, if e is an identity, $a^0 = e$.
- If a has an inverse, $n < 0$, $a^n = (a^{-1})^n$.
- These definitions lead to the rule: $a^m a^n = a^{m+n}\ \forall m, n \in \mathbb{Z}$.

Outline

Groups

Definition

A group is a set **G** with a composition law that is associative, has an identity element, and such that every element of **G** has an inverse.

A group **G** is abelian (or commutative) if the composition law is commutative.

The preferred notation for a group law is $(a, b) \to ab$ in general and $(a, b) \to a + b$ for abelian groups.

Abelian group

An abelian group is a set $\mathbf{G}$ with a composition law denoted '$+$'

$$\mathbf{G} \times \mathbf{G} \to \mathbf{G},$$

$$a, b \in \mathbf{G}, \quad a + b \in \mathbf{G}$$

satisfying the axioms

- associative $\forall a, b, c \in \mathbf{G}, \quad (a + b) + c = a + (b + c)$
- commutative $\forall a, b \in \mathbf{G}, \quad a + b = b + a$
- existence of O

$$\exists O \in \mathbf{G} \quad \text{such that } \forall a \quad a + O = a$$

- existence of inverses

$$\forall a \in \mathbf{G} \quad \exists b \in \mathbf{G} \quad \text{such that } a + b = O$$

This element is unique and denoted $-a$.

Examples of abelian groups

- $\mathbb{Z}^+$: the integers, with addition;
- $\mathbb{R}^+$: the real numbers, with addition;
- $\mathbb{R}^\times$: the nonzero real numbers, with multiplication;
- $\mathbb{C}$, $\mathbb{C}^\times$: analogous groups of complex numbers.

Let us get confused a bit!

A point worthy of discussion: Is it possible for the same set, say $\mathbb{R}$, to be an abelian group in more than one way? To show this, let us define a new addition of real numbers. We are going to call it 'O plus' $\oplus$:

$$a \oplus b := a + b - 1$$

Call this set $\mathbb{R}_{\oplus}$. It is easy to see that it is an abelian group [e.g. $(a \oplus b) \oplus c = a = b + c - 2$ so composition is associative] in which 0 is 1: $a \oplus 1 = a$!

Group of Rotations

Let C be the set of all complex numbers $a + bi$, with $a^2 + b^2 = 1$. Graphically this is just the unit circle centered at the origin of a plane. Ths set has the following properties:

- $a + bi \in C$, then $(a + bi)^{-1} \in C$. This because

$$(a + bi)^{-1} = (a - bi) \in C$$

- If $a + bi, c + di \in C$ then $(a + bi)(c + di) \in C$. This follows from $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$ and

$$(ac - bd)^2 + (ad + bc)^2 = (a^2 + b^2)(c^2 + d^2) = 1.$$

Each element of C can also be written

$$a + bi = e^{i\theta}$$

Cancellation Laws

Proposition

Let a, b, c be elements of the group $\mathbf{G}$*. If $ac = bc$, then $a = b$.*

Proof.

Multiply both sides of $ac = bc$ by c^{-1} on the right

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

$\square$

- Note the importance of the associative law.
- There is a similar cancellation law: If $ca = cb$, then $a = b$.

Groups: Main Examples

- For a positive integer n, let **G** be the set of invertible $n \times n$ matrices with entries in $\mathbb{R}$. If $\alpha, \beta \in$ **G**, multiplication yields a group structure on **G**. A notation for this group is $GL_n(\mathbb{R})$ and a characterization is

$$GL_n(\mathbb{R}) = \{\alpha \in \mathrm{Mat}_n(\mathbb{R}) : \det \alpha \neq 0\}$$

  Similar examples occur when other fields are used instead of $\mathbb{R}$.
- If **A** is a nonempty set, let **G** be the subset of bijective elements of $\mathrm{Maps}(\mathbf{A}, \mathbf{A})$. That is, the mapping $\alpha : \mathbf{A} \to \mathbf{A}$ is an element of **G** if α is one-one and onto. Such maps are also called of permutations of **A**. If **A** is a set of cardinality $n > 0$, **G** is denoted by S_n.

Examples

Consider the following two examples of groups:

- The set **G** of complex numbers of the form

$$\mathbf{G} = \{z = \cos(2\pi/4)n + i\sin(2\pi/4)n : n = 0, 1, 2, 3\}.$$

  Under multiplication **G** is a group with four elements.
- The set $\mathbb{Z}_4$ of residues of division by 4 of the integers, $\mathbb{Z}_4 = \{0, 1, 2, 3\}$. An addition can be given that defines a group law on $\mathbb{Z}_4$:

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

Comment: Except for the notation, it is hard to tell the difference between **G** and $\mathbb{Z}_4$. One feels tempted...

Isomorphisms

Definition

Let $\mathbf{G}$ and $\mathbf{G}'$ be two groups and $\varphi : \mathbf{G} \to \mathbf{G}'$ a mapping that is bijective. φ is a group isomorphism if

$$\varphi(ab) = \varphi(a)\varphi(b), \quad \forall a, b \in \mathbf{G}.$$

In the groups above, the function

$$\varphi : n \in \mathbb{Z}_4 \to \cos(2\pi/4)n + i\sin(2\pi/4)n \in \mathbf{G}$$

defines an isomorphism of the two groups. Another isomorphism would be the mapping

$$\varphi : n \in \mathbb{Z}_4 \to \cos(2\pi/4)3n + i\sin(2\pi/4)3n \in \mathbf{G}$$

Example

- Let **G** be the multiplicative group of positive real numbers, and **G**′ the additive group $\mathbb{R}$.
- The log function (in any base) gives a mapping

$$\log : \mathbf{G} \to \mathbf{G}'$$

with the property

$$\log(ab) = \log(a) + \log(b), \quad \forall a, b \in \mathbf{G}$$

is an isomorphism between the two groups.

- The corresponding exponential function gives the inverse isomorphism.

Outline

Subgroups

Definition

A subset **H** of a group **G** is a subgroup if it has the following three properties:

1. **Closure**: If $a, b \in \mathbf{H}$, then $ab \in \mathbf{H}$.
2. **Identity**: $1 \in \mathbf{H}$.
3. **Inverses**: If $a \in \mathbf{H}$, then $a^{-1} \in \mathbf{H}$.

Equivalently:

- **H** is a non-empty subset of **G** such that if $a, b \in \mathbf{H}$ then $ab^{-1} \in \mathbf{H}$.

It follows that **H** is a group for the restriction of the composition law of **G**.

Quaternions

Let $\mathbb{H}$ be the set of all 2×2 complex matrices of the form

$$\mathbf{A} = \begin{bmatrix} z_1 & z_2 \\ -\overline{z_2} & \overline{z_1} \end{bmatrix},$$

where z_1 and z_2 are complex numbers and $\overline{z_1}$ and $\overline{z_2}$ are their complex conjugates. These matrices satisfy:

- The product of two such matrices has the same format.
- If z_1 or z_2 is nonzero,

  $$\det \mathbf{A} = z_1\overline{z_1} + z_2\overline{z_2} \neq 0,$$

  so $\mathbf{A}$ is invertible, and $\mathbf{A}^{-1}$ has the same format.
- Thus the set of nonzero matrices of $\mathbb{H}$ form a group. ($\mathbb{H}$ is even grander!)

Example

The matrices of **H**

$$1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{I} = \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix}$$

$$\mathbf{J} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

satisfy

$$\mathbf{IJ} = -\mathbf{JI} = \mathbf{K}, \quad \mathbf{JK} = -\mathbf{KJ} = \mathbf{I}, \quad \mathbf{KI} = -\mathbf{IK} = \mathbf{J}$$

so the 8 elements $\{\pm 1, \pm\mathbf{I}, \pm\mathbf{J}, \pm\mathbf{K}\}$ form a subgroup.

The subgroups of $\mathbb{Z}^+$

Theorem

If a is an integer, the set

$$\mathbf{H} = \{na : n \in \mathbb{Z}\}$$

is a subgroup of $\mathbb{Z}$. It is usually denoted by $\mathbb{Z}a$.

Conversely, if $\mathbf{H}$ is a subgroup of $\mathbb{Z}$, then there is an integer $a \geq 0$ such that $\mathbf{H}$ consists of all multiples of a.

Proof. It is clear that $\mathbb{Z}a$ is a subgroup: If $na, ma \in \mathbf{H}$,

$$na \pm ma = (n \pm m)a \in \mathbf{H}.$$

Converse

For the converse, if $\mathbf{H}$ is the subgroup $\{0\}$, then $\mathbf{H} = \mathbb{Z}0$. If not, $\mathbf{H}$ contains positive integers: if $b \in \mathbf{H}$ and $b < 0$, $-b \in \mathbf{H}$.
Let a be the smallest nonzero positive integer in $\mathbf{H}$. We claim that $\mathbf{H} = \mathbb{Z}a$. For $b \in \mathbf{H}$, by the Euclidean algorithm, there is a relation

$$b = qa + r, \quad 0 \leq r < a.$$

If $r = 0$, $b \in \mathbb{Z}a$. If not, $r = b - qa \in \mathbf{H}$, which is a contradiction since $0 < r < a$.

Application

Let a and b be two integers. The set

$$\mathbf{H} = \{ma + nb : m, n \in \mathbb{Z}\}$$

has the required properties of a subgroup of $\mathbb{Z}$: closure, identity, negatives.
According to the Proposition, $\mathbf{H} = \mathbb{Z}c$ for some integer c. Let us examine the properties of c:

1. Since $a, b \in \mathbf{H}$, a and b are divisible by c.
2. Any integer d that divides a and b, will divide a combination $ma + nb$, in particular will divide c.
3. This shows that $c = \gcd(a, b)$.
4. As extra we obtain that the gcd of a and b is a linear combination $ma + nb$ for appropriate $m, n \in \mathbb{Z}$.

Cyclic Groups

One of the most natural ways to create groups is the following: Let **G** a group. Let x be a fixed element of **G** and set

$$\langle x \rangle = \{x^n : n \in \mathbb{Z}\}.$$

This is the set of all powers of x. Recall that $x^0 = 1$ and if $n < 0$ $x^n = (x^{-1})^{-n}$. $\langle x \rangle$ is clearly a subgroup of **G**: has closure, identity and inverse properties.

Definition

$\langle x \rangle$ is called a cyclic group and x is called a generator.

$\mathbb{Z}$ **is cyclic**, and ± 1 are its generators: any integer is a multiple of any of them.

Properties of Cyclic Groups

The following is elementary but very important:

Proposition

Let $\mathbf{G} = \langle x \rangle$ *be a cyclic group.*

1. *If* $\mathbf{G}$ *is infinite then* $\mathbf{G}$ *is isomorphic to* $\mathbb{Z}$.
2. *Every subgroup* $\mathbf{H}$ *of a cyclic group is cyclic.*
3. *Two cyclic groups of the same cardinality are isomorphic.*

Proof

Proof. One volunteer please!

Order of a group

Definition

The order of a group $\mathbf{G}$ is its cardinality: $|\mathbf{G}|$.
If $a \in \mathbf{G}$, its *order* is the smallest integer n(or infinity if n does not exist) such that $a^n = 1$. Thus the order of a is the order of the cyclic subgroup generated by a: $|\langle a \rangle|$

Example: $|S_n| = n!$

Homomorphisms

Definition

Let $\mathbf{G}$ and $\mathbf{G}'$ be groups. A homomorphism $\varphi : \mathbf{G} \to \mathbf{G}'$ is any mapping satisfying rule

$$\varphi(ab) = \varphi(a)\varphi(b), \quad \forall a, b \in \mathbf{G}.$$

Example

- the determinant function $\det : GL_n(\mathbb{R}) \to \mathbb{R}^\times$;
- the logarithm $\log : \mathbb{R}_+^\times \to \mathbb{R}$;
- the exponential...
- the mapping $\varphi : \mathbb{Z} \to \mathbf{G}$, $\varphi(n) = a^n$, where a is a fixed element of $\mathbf{G}$;
- the inclusion map $i : \mathbf{H} \to \mathbf{G}$ where $\mathbf{H}$ is a subgroup of $\mathbf{G}$

Conjugation

Definition

Let $\mathbf{G}$ be a group and $a \in \mathbf{G}$. The conjugation defined by a is the mapping $\varphi : \mathbf{G} \to \mathbf{G}$

$$\varphi(b) = aba^{-1}, \quad b \in \mathbf{G}.$$

Proposition

φ is a homomorphism.

Proof.

$$a(bc)a^{-1} = (aba^{-1})(aca^{-1}).$$

$\square$

Properties of Homomorphisms

Proposition

If $\varphi : \mathbf{G} \to \mathbf{G}'$ is a group homomorphism, then it carries the identity to the identity, and inverses to inverses. In other words, $\varphi(1) = 1'$ and $\varphi(a^{-1}) = (\varphi(a))^{-1}$.

Proof. (a) $\varphi(1) = \varphi(1 \cdot 1) = \varphi(1)\varphi(1)$, so by cancellation $\varphi(1) = 1'$.

(b) $1 = a \cdot a^{-1} \to \varphi(a)\varphi(a^{-1}) = \varphi(1) = 1'$. Thus $\varphi(a^{-1})$ is the inverse of $\varphi(a)$.

Properties of Homomorphisms

Definition

Let $\varphi : \mathbf{G} \to \mathbf{G}'$ be a group homomorphism.

1. The kernel of φ is the set

$$\mathbf{H} = \{a \in \mathbf{G} : \varphi(a) = 1'\}.$$

2. The range, or image of φ is the set

$$\mathbf{L} = \{\varphi(x) : x \in \mathbf{G}\}.$$

Proposition

Let $\varphi : \mathbf{G} \to \mathbf{G}'$ be a group homomorphism. Then the kernel of φ is a subgroup of $\mathbf{G}$ and its image is a subgroup of $\mathbf{G}'$.

Proof. A volunteer please!

Another volunteer to discuss:

$$\varphi : \mathbb{C}^\times \to \mathbb{C}^\times, \quad \varphi(z) = z^4$$

Normal Subgroups

A super important type of subgroup are normal sybgroups. They were [introduced and] extensively used by Galois in his analysis of the solvability of equations by radicals.

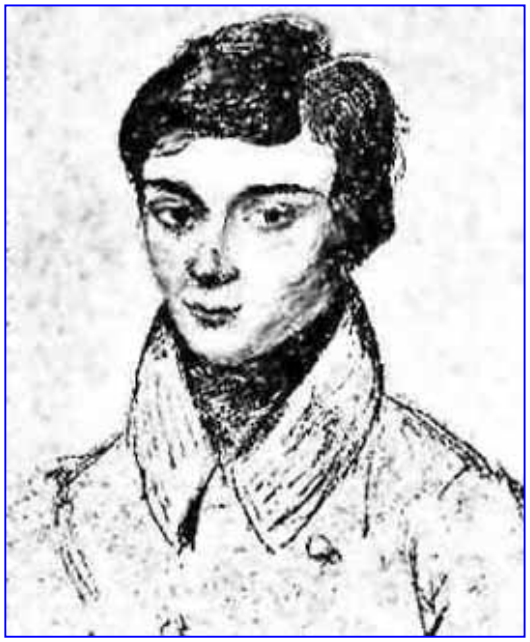
Definition

A subgroup **H** of a group **G** is a normal subgroup if it has the following property: For every $a \in \mathbf{H}$ and every $b \in \mathbf{G}$, the conjugate bab^{-1} is in **H**.

Evariste Galois (1811-1832)

Galois Portraits

Evariste Galois



A drawing done in 1848 from memory by Evariste's brother

David Hilbert (1862-1943)

David Hilbert

David Hilbert
(1862 - 1943)
Mathematician
Algebraist
Topologist
Geometrist
Number Theorist
Physicist
Analyst
Philosopher
Genius
And modest too...

"Physics is much too hard for physicists." - Hilbert, 1912

Examples

- If $\mathbf{G}$ is abelian, then every subgroup is normal.
- If $\varphi : \mathbf{G} \to \mathbf{G}'$ is a group homomorphism, then its kernel $\mathbf{H}$ is normal: If $\varphi(a) = 1'$ then for any $b \in \mathbf{G}$,

$$\varphi(bab^{-1}) = \varphi(b)\varphi(a)\varphi(b^{-1}) = \varphi(b) \cdot 1' \cdot \varphi(b)^{-1} = 1'.$$

- The subgroup $SL_n(\mathbb{R}) \subset GL_n(\mathbb{R})$ real matrices of determinant 1 is normal: It is the kernel of $\det : GL_n(\mathbb{R}) \to \mathbb{R}^\times$.
- If $\mathbf{H}$ is a subgroup of a group $\mathbf{G}$ and $[\mathbf{G} : \mathbf{H}] = 2$, then $\mathbf{H}$ is normal: $\mathbf{G} = \mathbf{H} \cup a\mathbf{H} = \mathbf{H} \cup \mathbf{H}a$, thus $a\mathbf{H} = \mathbf{H}a$.

Normalizer of a Subgroup

Let **H** be a subgroup of a group **G**. The set

$$N(\mathbf{H}) = \{x \in \mathbf{G} : x\mathbf{H} = \mathbf{H}x\}$$

is a subgroup. **H** is a normal subgroup of $N(\mathbf{H})$.

Definition

$N(\mathbf{H})$ is the normalizer of **H**.

Given a subgroup **H** of a group **G**, $N(\mathbf{H})$ is the largest subgroup **K** such that **H** is a normal subgroup of...

Outline

Quick Review: Vector Spaces

A vector space is a structured set put together from an abelian group **V** and a field **F**. It is helpful to keep in mind the following examples.

Let n be a non-negative integer. $\mathbb{R}^n$: the set of all n-tuples of real numbers, with 2 compositions

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} v_1 + u_1 \\ v_2 + u_2 \\ \vdots \\ v_n + u_n \end{bmatrix}$$

For $c \in \mathbb{R}$,

$$c \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix}$$

Another example is the set of polynomials in one indeterminate over the field $\mathbf{F}$: $\mathbf{F}[x]$ is the set of polynomials

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \quad a_i \in \mathbf{F}$$

Addition is given by

$$(a_n x^n + \cdots + a_1 x + a_0) + (b_m x^m + \cdots + b_1 x + b_0) = \sum_i (a_i + b_i) x^i$$

and scalar multiplication

$$cf(x) = ca_n x^n + ca_{n-1} x^{n-1} + \cdots + ca_1 x + ca_0$$

Related examples are the subsets $\mathbb{P}_n(x)$ of polynomials of degree at most n.

The set of solutions of the differential equation

$$y^{(3)} - 7y'' + 14y' - 8y = 0$$

is also a vector space over $\mathbb{R}$. It is a consequence of the fact [principle of superposition] that if $y_1(x)$ and $y_2(x)$ are solutions then for $a, b \in \mathbb{R}$

$$ay_1(x) + by_2(x)$$

is also a solution. From Calc 252, it will follow that any solution is a combination

$$ae^x + be^{2x} + ce^{4x}$$

Formally, a vector space over a field $\mathbf{F}$ is an abelian group $\mathbf{V}$ admitting a (scalar) multiplication

$$\mathbf{F} \times \mathbf{V} \to \mathbf{V}, \quad c \times u \mapsto cu \in \mathbf{V}$$

with the following properties:

- For $c, d \in \mathbf{F}$, $u \in \mathbf{V}$, $(cd)u = c(du)$
- For $u \in \mathbf{V}$, $1u = u$
- For $c, d \in \mathbf{F}$, $u \in \mathbf{V}$, $(c + d)u = cu + du$
- For $c \in \mathbf{F}$, $u, v \in \mathbf{V}$, $c(u + v) = cu + cv$

We can now define vectors: the elements of a vector space.

Theorem (First Theorem)

For $u, O \in \mathbf{V}$, $0, c \in \mathbf{F}$

$$0u = O, \quad cO = O, \quad (-c)u = -(cu)$$

Proof. For the first claim, observe

$$0u = (0 + 0)u = 0u + 0u,$$

so

$$0u = O$$

Similarly for the other claims. $\square$

There are many vector spaces derived from those mentioned already. We give a very general method to form new vector spaces. Let $\mathbf{V}$ and $\mathbf{W}$ be vector spaces over the field $\mathbf{F}$ and let $\mathbf{V} \times \mathbf{W}$ be the set of all ordered pairs (v, w), $v \in \mathbf{V}$, $w \in \mathbf{W}$. If we define an addition and a scalar multiplication by

$$\begin{aligned} (v_1, w_1) + (v_2, w_2) &:= (v_1 + v_2, w_1 + w_2) \\ c(v, w) &:= (cv, cw), \end{aligned}$$

we make $\mathbf{V} \times \mathbf{W}$ into a vector space. It is easy to verify all the requirements. This is the method used to obtain the vector spaces of tuples $\mathbf{F}^2 = \mathbf{F} \times \mathbf{F}$, $\mathbf{F}^3 = \mathbf{F}^2 \times \mathbf{F}$, and so on.

Functions on Vector Spaces

Let **V** and **W** be two vector spaces over the field **F**. What are the functions like between these spaces?:

$$\mathbf{T} : \mathbf{V} \to \mathbf{W}.$$

V is called the source, and **W** the target of the function. For example, suppose $\mathbf{V} = \mathbf{W} = \mathbf{F}^2$. Then **T** takes for input pairs $v = (x_1, x_2)$, and outputs pairs $\mathbf{T}(v) = (y_1, y_2)$:

$$(x_1, x_2) \to \boxed{\quad \mathbf{T} \quad} \to (y_1, y_2) = (\mathbf{f}_1(x_1, x_2), \mathbf{f}_2(x_1, x_2))$$

It can be very varied since functions of two variables come in many flavors.

We will be looking at certain type of functions illustrated by the following examples.

- Let **V** be the vector space of all real valued functions with derivatives in $[-1, 1]$, and let **W** be the vector space of real valued functions on $[-1, 1]$. Define

$$\mathbf{T}(f(t)) = f'(t),$$

or

$$\mathbf{L}(f) = f'' - f.$$

- Here are two other functions

$$\mathbf{T}(f) = \int_{-1}^{1} f(t)dt, \quad \mathbf{T} : \mathbf{V} \to \mathbb{R}$$

$$\mathbf{L}(f) = \int_{-1}^{t} f(t)dt, \quad \mathbf{T} : \mathbf{V} \to \mathbf{W}$$

- $\mathbf{T} : \mathbb{R}^2 \to \mathbb{R}^2$
$$\mathbf{T}(x, y) = (y, x)$$
This is reflection about the [main] diagonal.
- For α fixed,
$$\mathbf{T}(x, y) = (x \cos \alpha + y \sin \alpha, -x \sin \alpha + y \cos \alpha)$$
This is a rotation in the plane by α degrees.
- $\mathbf{T} : \mathbb{R}^3 \to \mathbb{R}^2$
$$\mathbf{T}(x, y, z) = (x, y)$$
This is projection on the xy-plane.

All these functions share the following property:

Definition

A function $\mathbf{T} : \mathbf{V} \to \mathbf{W}$ is a linear transformation, or linear operator, if it satisfies:

(i) For any $v_1, v_2 \in \mathbf{V}$,

$$\mathbf{T}(v_1 + v_2) = \mathbf{T}(v_1) + \mathbf{T}(v_2)$$

[$\mathbf{T}$ is additive, that is takes sums to sums]

(ii) For any $v \in \mathbf{V}$ and $c \in \mathbf{F}$,

$$\mathbf{T}(cv) = c\mathbf{T}(v)$$

[$\mathbf{T}$ commutes with scaling]

We can put these two properties together:
A function can be viewed as a factory processing inputs into outputs

$$\text{input} \longrightarrow \boxed{\quad \mathbf{f} \quad} \longrightarrow \text{output}$$

One key property of a linear box is that it can be reverse engineered.

Proposition

Let $\mathbf{T} : \mathbf{V} \to \mathbf{W}$ *be a linear transformation of vector spaces. If* $v_1, \ldots, v_n \in \mathbf{V}$ *and* $c_1, \ldots, c_n \in \mathbf{F}$*, then*

$$\mathbf{T}(\sum_{i=1}^{n} c_i v_i) = \sum_{i=1}^{n} c_i \mathbf{T}(v_i).$$

[That is, $\mathbf{T}$ *commutes with linear combinations.]*

Proof.

It uses the conditions (i) and (ii) of the definition and induction: Apply $\mathbf{T}$ to

$$\sum_{i=1}^{n} c_i v_i = (\sum_{i=1}^{n-1} c_i v_i) + c_n v_n$$

and iterate. □

Recipe for linear transformations

Let **V** be a vector space with a basis $v_1, v_2, \ldots$. If **W** is a vector space, for each v_i choose $w_i \in \mathbf{W}$ [the w_i do not need to be linearly independent].

Proposition

The assignment

$$\sum_i x_i v_i \mapsto \sum_i x_i w_i$$

defines a linear transformation from **V** *to* **W**.

One quick way to build a L.T. between spaces of tuples is the following. Let $\mathbf{A}$ be an $m \times n$ matrix with entries in the field $\mathbf{F}$. For a n-tuple

$$v = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

define the function $\mathbf{L_A} : \mathbf{F}^n \to \mathbf{F}^m$

$$\mathbf{L_A}(v) = \mathbf{A} \cdot v.$$

Since multiplication of matrices is distributive and commutes with scaling, $\mathbf{L_A}$ is a L.T.

Let $\mathbf{M}_2(\mathbf{F})$ be the vector space of all 2-by-2 matrices over the field $\mathbf{F}$. Fix a matrix, say, $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$ and define the function

$$\mathbf{B} \to \mathbf{T}(\mathbf{B}) = \mathbf{AB}.$$

$\mathbf{T}$ satisfies:

$$\begin{aligned} \mathbf{T}(\mathbf{B}_1 + \mathbf{B}_2) &= \mathbf{A}(\mathbf{B}_1 + \mathbf{B}_2) = \mathbf{AB}_1 + \mathbf{AB}_2 \\ &= \mathbf{T}(\mathbf{B}_1) + \mathbf{T}(\mathbf{B}_2) \\ \mathbf{T}(c\mathbf{B}) &= \mathbf{A}(c\mathbf{B}) = c\mathbf{AB} = c\mathbf{T}(B). \end{aligned}$$

Point: Lots of freedom to create linear transformations.

There are several subsets associated to a linear transformation $\mathbf{T} : \mathbf{V} \to \mathbf{W}$:

- The Nullspace or Kernel of $\mathbf{T}$ is the subset

$$N(\mathbf{T}) = \{v \in \mathbf{V} \mid \mathbf{T}(v) = O.\}$$

  [The vectors mapped to O]

- The Range or Image of $\mathbf{T}$ is the subset

$$R(\mathbf{T}) = \{w \in \mathbf{W} \mid w = \mathbf{T}(v), \quad v \in \mathbf{V}.\}$$

Examples

If **T** is the linear transformation

$$f \mapsto f'' - f$$

defined earlier, its nullspace consists of the solutions of $y'' - y = 0$, that is the linear combinations

$$ae^x + be^{-x}, \quad a, b \in \mathbb{R}.$$

Proposition

The Nullspace and the Range of a linear transformation $\mathbf{T} : \mathbf{V} \to \mathbf{W}$ *are subspaces of* $\mathbf{V}$ *and* $\mathbf{W}$ *respectively.*

Proof.

Let us apply the subspace test to $N(\mathbf{T})$. Suppose $v_1, v_2 \in N(\mathbf{T})$. Then for any scalars c_1, c_2,

$$\mathbf{T}(c_1 v_1 + c_2 v_2) = c_1 \mathbf{T}(v_1) + c_2 \mathbf{T}(v_2) = c_1 O + c_2 O = O.$$

So the linear combination belongs to the Nullspace.
We leave for you the other proof. □

The dimension of $N(\mathbf{T})$ is called the nullity and the dimension of $R(\mathbf{T})$ is called the rank of $\mathbf{T}$.

Dimension Formula

Theorem

Let $\mathbf{T} : \mathbf{V} \to \mathbf{W}$ *be a linear transformation of finite dimensional vector spaces. Then*

$$\dim N(\mathbf{T}) + \dim R(\mathbf{T}) = \dim \mathbf{V}.$$

That is, nullity $+$ *rank* $= \dim \mathbf{V}$.

Proof. Suppose $v_1, \ldots, v_n$ is a basis of $\mathbf{V}$, and $u_1, \ldots, u_r$ is a basis of $N(\mathbf{T})$. We are going to show that $R(\mathbf{T})$ has basis with $n - r$ elements.
Recall that $\mathbf{T}(\sum_{i=1}^{n} c_i v_i) = \sum_{i=1}^{n} c_i \mathbf{T}(v_i)$, $R(\mathbf{T})$ is spanned by

$$\mathbf{T}(v_1), \mathbf{T}(v_2), \ldots, \mathbf{T}(v_j), \ldots, \mathbf{T}(v_n).$$

Out of this list we are going to pick a basis for $R(\mathbf{T})$.

We scan the list and delete the vectors [red] that can be written as linear combination of the preceding vectors

$$\mathbf{T}(v_1), \mathbf{T}(v_2), \ldots, \mathbf{T}(v_j), \ldots, \mathbf{T}(v_n).$$

For convenience of notation we assume we are left with the first s vectors

$$\mathbf{T}(v_1), \mathbf{T}(v_2), \ldots, \mathbf{T}(v_s).$$

Claim: $u_1, \ldots, u_r, v_1, \ldots, v_s$ is a basis of $\mathbf{V}$.
Once we have shown this we be done since all bases of $\mathbf{V}$ have n elements. Let us check.

Claim 1: $u_1, \ldots, u_r, v_1, \ldots, v_s$ spans $\mathbf{V}$

If $v \in \mathbf{V}$, $\mathbf{T}(v) = \sum_{i=1}^{s} a_i \mathbf{T}(v_i)$, that is

$$\mathbf{T}(v - \sum_{i=1}^{s} a_i v_i) = 0$$

that is $v - \sum_{i=1}^{s} a_i v_i$ belongs to the nullspace so

$$v - \sum_{i=1}^{s} a_i v_i = \sum_{j=1}^{r} b_j u_j.$$

Claim 2: $u_1, \ldots, u_r, v_1, \ldots, v_s$ are linearly independent.

If $\sum b_j u_j + \sum a_i v_i = O$, applying $\mathbf{T}$ we get $\sum a_i \mathbf{T}(v_i) = 0$ since $\mathbf{T}(u_j) = 0$. But the $\mathbf{T}(v_i)$ are linearly independent [they form a basis of $R(\mathbf{T})$] so $a_i = 0$. We have left $\sum b_j u_j = 0$, which implies $b_j = 0$ since the u_j form a basis of $N(\mathbf{V})$

Let us recall some general properties of a function $\mathbf{f} : \mathbb{X} \to \mathbf{Y}$

- $\mathbf{f}$ is one-one if $\mathbf{f}(x_1) = \mathbf{f}(x_2)$ implies $x_1 = x_2$. One also says that $\mathbf{f}$ is injective. If $\mathbf{f}$ is a linear transformation, $\mathbf{f}(x_1) = \mathbf{f}(x_2)$ means $\mathbf{f}(x_1 - x_2) = O$ so $\mathbf{f}$ is one-one if and only if the nullspace is (O).
- $\mathbf{f}$ is onto if its image is $\mathbf{Y}$. One also says that $\mathbf{f}$ is surjective.
- $\mathbf{f}$ is an isomorphism or invertible when it is both.

Here are some consequences of the dimension formula applied to a linear transformation $\mathbf{T} : \mathbf{V} \to \mathbf{W}$

- If $\dim \mathbf{V} > \dim \mathbf{W}$, then $\mathbf{T}$ is not one-one
- If $\dim \mathbf{V} < \dim \mathbf{W}$, then $\mathbf{T}$ is not onto.
- If $\dim \mathbf{V} = \dim \mathbf{W}$, then $\mathbf{T}$ is an isomorphism iff its nullspace is O, or iff is onto.

Outline

We shall now discuss how to represent some [look at the caveat] linear transformations $\mathbf{T} : \mathbf{V} \to \mathbf{W}$ by matrices. It is a process akin to representing vectors by coordinates. Recall that if $v \in \mathbf{V}$ and $\mathcal{B} = v_1, \ldots, v_n$ is a basis of $\mathbf{V}$, we have a unique expression

$$v = x_1 v_1 + \cdots + x_n v_n.$$

We say that the x_i are the coordinates of v with respect to $\mathcal{B}$. We write as

$$[v]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

If $\mathcal{C} = w_1, \ldots, w_m$ we would like to find the coordinates of $\mathbf{T}(v)$ in the basis $\mathcal{C}$

$$[\mathbf{T}(v)]_{\mathcal{C}} = \begin{bmatrix} ? \end{bmatrix}.$$

In other words, if $v = x_1 v_1 + \cdots + x_n v_n$,

$$\mathbf{T}(v) = y_1 w_1 + \cdots + y_m w_m,$$

we want to describe the y_i in terms of the x_j. The process will be called a matrix representation. It comes about as follows:

$$\sum y_i w_i = T(\sum x_j v_j) = \sum x_j \mathbf{T}(v_j)$$

Thus if we have the coordinates of the $\mathbf{T}(v_j)$,

$$\mathbf{T}(v_j) = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix}$$

we have

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \sum x_j \begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix}$$

More pictorially

$$[\mathbf{T}(v)]_{\mathcal{C}} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = [\mathbf{T}]_{\mathcal{B}}^{\mathcal{C}} \cdot [v]_{\mathcal{B}}$$

The $n \times m$ matrix

$$[\mathbf{T}]_{\mathcal{B}}^{\mathcal{C}}$$

is called the matrix representation of $\mathbf{T}$ relative to the bases $\mathcal{B}$ of $\mathbf{V}$ and $\mathcal{C}$ of $\mathbf{W}$.

Quickly: Once bases $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$ have been chosen, $\mathbf{T}$ is represented by

$$\left[\ a_{ij}\ \right]$$

where the entries come from

$$\mathbf{T}(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Example

Recall the transpose operation on a square matrix $\mathbf{A}$: if a_{ij} is the (i,j)-entry of $\mathbf{A}$, the (i,j)-entry of $\mathbf{A}^t$ is a_{ji}. This is a linear transformation $\mathbf{T}$ on the space $\mathbf{M}_n(\mathbf{F})$:

$$(\mathbf{A}+\mathbf{B})^t = \mathbf{A}^t + \mathbf{B}^t, \quad (c\mathbf{A})^t = c\mathbf{A}^t.$$

Let us find its matrix representation on $\mathbf{M}_2(\mathbf{F})$. This space has the basis

$$v_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, v_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, v_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Since

$$\mathbf{T}(v_1) = v_1, \quad \mathbf{T}(v_2) = v_3, \quad \mathbf{T}(v_3) = v_2, \quad \mathbf{T}(v_4) = v_4,$$

the matrix representation of transposing is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Let $\mathbb{R}_3[x]$ be the space of real polynomials of degree at most 3 and **T** the differentiation operator.

A basis here are the polynomials $1, x, x^2, x^3$. The corresponding matrix representation is

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Exercise 2: Suppose a linear transformation $\mathbf{T} : \mathbb{R}^3 \to \mathbb{R}^2$ satisfies

$$\mathbf{T}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{T}\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \mathbf{T}\left(\begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

(a) Show that the three vectors of $\mathbb{R}^3$ are linearly independent.
(b) Find the nullspace of this linear transformation.
(c) Find $\mathbf{T}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right)$.

Let us solve an exercise that usually gets a shaky argument. Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ matrices such that $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}$.

Claim: $\mathbf{B} \cdot \mathbf{A} = \mathbf{I}$. [The question arises because matrix multiplication is not commutative.] To argue we consider the L.T.s $\mathbf{L_A}$ and $\mathbf{L_B}$ associated to $\mathbf{A}$ and $\mathbf{B}$.

$\mathbf{A} \cdot \mathbf{B} = \mathbf{I}$ implies that

$$\mathbf{L_A} \circ \mathbf{L_B} = \mathbf{I},$$

from which it follows that $\mathbf{L_B}$ is one-one, and therefore it is invertible, so

$$\mathbf{L_A} \circ \mathbf{L_B} = \mathbf{L_B} \circ \mathbf{L_A} = \mathbf{I}.$$

Thus $\mathbf{B} \cdot \mathbf{A} = \mathbf{I}$.

Blocks

Suppose $\mathbf{T}$ is a L.T. of vector space $\mathbf{V}$ with a basis $\mathcal{A} = v_1, \ldots, v_r, v_{r+1}, \ldots, v_n$. Suppose $\mathbf{T}(v_i)$ for $i \leq r$, is a linear combination of the first r basis vectors, and $\mathbf{T}(v_i)$ for $i > r$, is a linear combination of the last $n - r$ basis vectors.
Claim: The matrix representation has the block format

$$[\mathbf{T}]_{\mathcal{A}} = \begin{bmatrix} \boxed{r \times r} & O \\ O & \boxed{(n-r) \times (n-r)} \end{bmatrix}$$

This can be refined to more than two blocks. The extreme case is when all blocks are 1×1. The representation is then said to be diagonal. If and when this happens is a major theme of Linear Algebra.

Addition of linear transformations

We are now going to combine linear transformations in various ways.
Let **T** and **U** be two linear transformations of source **V** and target **W**. Consider the operations,

$$\begin{aligned}(\mathbf{T}+\mathbf{U})(v) &:= \mathbf{T}(v)+\mathbf{U}(v)\\ (c\mathbf{T})(v) &:= c\mathbf{T}(v).\end{aligned}$$

Clearly they define [write the reasons] a vector space on the set $\mathcal{L}(\mathbf{V},\mathbf{W})$ of all such linear transformations.

Theorem

If **V** *has dimension n and* **W** *has dimension m, then*

$$\dim \mathcal{L}(\mathbf{V},\mathbf{W}) = m \cdot n.$$

We are going to build a basis for this space. Let $\mathcal{B} = v_1, \ldots, v_n$ be a basis of $\mathbf{V}$ and $\mathcal{C} = w_1, \ldots, v_m$ be a basis of $\mathbf{W}$. Using the basic recipe, define the linear transformation

$$\mathbf{E}_{ij}(v_k) = \begin{cases} 0, & k \neq i \\ w_j, & k = i \end{cases}$$

There are $m \cdot n$ such [elementary] linear transformations.

Proposition

The $\mathbf{E}_{ij}$ *are linearly independent. [Which also follows from the next assertion.] If* $\mathbf{T}$ *is a linear transformation and*

$$\mathbf{T}(v_j) = \sum_i a_{ij} w_j,$$

then

$$\mathbf{T} = \sum_{i,j} a_{ij} \mathbf{E}_{ij}.$$

Proof. Try yourself or look up in book.

Composition of linear transformations

There is another way to combine certain linear transformations. Consider composition of functions

$$\mathbf{V} \xrightarrow{\mathbf{T}} \mathbf{W} \xrightarrow{\mathbf{U}} \mathbf{Z},$$

$$(\mathbf{U} \circ \mathbf{T})(v) := \mathbf{U}(\mathbf{T}(v))$$

Proposition

With $\mathbf{T}$ *and* $\mathbf{U}$ *as above,* $\mathbf{U} \circ \mathbf{T}$ *is a linear transformation from* $\mathbf{V}$ *to* $\mathbf{Z}$*.*

Proof.

Let us check the basic requirements:

$$\begin{aligned}\mathbf{U} \circ \mathbf{T}(v_1 + v_2) &:= \mathbf{U}(\mathbf{T}(v_1 + v_2)) = \mathbf{U}(\mathbf{T}(v_1) + \mathbf{T}(v_2)) \\ &= \mathbf{U}(\mathbf{T}(v_1)) + \mathbf{U}(\mathbf{T}(v_2)) \\ &= \mathbf{U} \circ \mathbf{T}(v_1) + \mathbf{U} \circ \mathbf{T}(v_2).\end{aligned}$$

It shows composition is additive.

$$\begin{aligned}\mathbf{U} \circ \mathbf{T}(cv) &:= \mathbf{U}(\mathbf{T}(cv)) = \mathbf{U}(c\mathbf{T}(v)) \\ &= c\mathbf{U}(\mathbf{T}(v)) = c(\mathbf{U} \circ \mathbf{T})(v).\end{aligned}$$

It shows the scaling property. □

Now we are going to explain where multiplication of matrices comes from and why it is associative.Suppose we have a composition of L.T.'s [linear transformations]

$$\mathbf{V} \xrightarrow{\mathbf{T}} \mathbf{W} \xrightarrow{\mathbf{U}} \mathbf{Z},$$

and that we have chosen bases $\mathcal{B}$, $\mathcal{C}$ $\mathcal{D}$, so that we have matrix representations

$$[\mathbf{T}]_{\mathcal{B}}^{\mathcal{C}}, \quad [\mathbf{U}]_{\mathcal{C}}^{\mathcal{D}}.$$

Theorem

The matrix representation of the composition $\mathbf{U} \circ \mathbf{T}$ *is*

$$[\mathbf{U} \circ \mathbf{T}]_{\mathcal{B}}^{\mathcal{D}} = [\mathbf{U}]_{\mathcal{C}}^{\mathcal{D}} \circ [\mathbf{T}]_{\mathcal{B}}^{\mathcal{C}}.$$

To prove this we pick the bases $\mathcal{B} = v_1, \ldots, v_n$, $\mathcal{C} = u_1, \ldots, u_m$, $\mathcal{D} = w_1, \ldots, w_p$, and look for the coefficient c_{ji} of w_i in the expression of $(\mathbf{U} \circ \mathbf{T})(v_j)$:

$$\begin{aligned}
(\mathbf{U} \circ \mathbf{T})(v_j) &= \mathbf{U}(\mathbf{T}(v_j)) = \mathbf{U}(\sum_{k=1}^{m} a_{kj} u_k) \\
&= \sum_{k=1}^{m} a_{kj} \mathbf{U}(u_k) = \sum_{k=1}^{m} a_{kj} (\sum_{\ell=1}^{p} b_{\ell k} w_\ell) \\
&= \sum_{\ell=1}^{p} (\sum_{k=1}^{m} b_{\ell k} a_{kj}) w_\ell
\end{aligned}$$

This gives

$$c_{ij} = \sum_{k=1}^{m} b_{ik} a_{kj},$$

the usual row by column rule of multiplication.

There is a consequence that is tedious to verify directly, that the product of matrices is associative:

This follows from the tautology of the composition of functions

$$(\mathbf{A} \circ \mathbf{B}) \circ \mathbf{C} = \mathbf{A}(\mathbf{B}(\mathbf{C})) = \mathbf{A} \circ (\mathbf{B} \circ \mathbf{C})$$

and the theorem above.

Invertible linear transformations

Let

$$\mathbf{T} : \mathbf{V} \to \mathbf{W}$$

be a L.T. that is one-one and onto. This means that for any $w \in \mathbf{W}$ there is a unique $v \in \mathbf{V}$ such that $\mathbf{T}(v) = w$. This gives rise to a function

$$\mathbf{U} : \mathbf{W} \to \mathbf{V}, \quad \mathbf{U}(w) = v \quad \text{iff} \quad \mathbf{T}(v) = w.$$

$\mathbf{U}$ is the inverse function of $\mathbf{T}$:

$$\mathbf{U} \circ \mathbf{T} = \mathbf{I_V} \quad \text{the identity of } \mathbf{V}.$$

One also checks

$$\mathbf{T} \circ \mathbf{U} = \mathbf{I_W} \quad \text{the identity of } \mathbf{W}.$$

Notation: $\mathbf{U} = \mathbf{T}^{-1}$.

Proposition

If $\mathbf{T}$ *is a L.T. then* $\mathbf{U}$ *is also a L.T.*

Proof.

Let $w_1, w_2 \in \mathbf{W}$. Pick $v_1, v_2 \in \mathbf{V}$ so that $\mathbf{T}(v_1) = w_1$ and $\mathbf{T}(v_2) = w_2$. Since $\mathbf{T}$ is a L.T.,

$$\mathbf{T}(v_1 + v_2) = \mathbf{T}(v_1) + \mathbf{T}(v_2) = w_1 + w_2.$$

By the definition of $\mathbf{U}$,

$$\mathbf{U}(w_1 + w_2) = v_1 + v_2 = \mathbf{U}(w_1) + \mathbf{U}(w_2),$$

so $\mathbf{U}$ is additive. The scaling property is proved in the same way. $\square$

If $\mathbf{T} : \mathbf{V} \to \mathbf{W}$ is an invertible L.T., choosing bases $\mathcal{B}$ and $\mathcal{C}$ for the two spaces:

Proposition

The matrix representations of $\mathbf{T}$ *and* $\mathbf{T}^{-1}$ *are related as follows*

$$[\mathbf{T}^{-1}]_{\mathcal{C}}^{\mathcal{B}} = ([\mathbf{T}]_{\mathcal{B}}^{\mathcal{C}})^{-1}.$$

Proof.

This follows from the equalities

$$\mathbf{T}^{-1} \circ \mathbf{T} = \mathbf{I}_{\mathbf{V}}, \quad \mathbf{T} \circ \mathbf{T}^{-1} = \mathbf{I}_{\mathbf{W}}$$

and a previous result asserting that the matrix representation of a composition of two L.T. is the product of the matrices. □

If $\mathbf{T}$ is invertible, we also say that it is an isomorphism, and that $\mathbf{V}$ and $\mathbf{W}$ are isomorphic vector spaces. For this to happens it requires that $\dim \mathbf{V} = \dim \mathbf{W}$.

Example: Let $\mathbb{P}_4[x]$ be the space of polynomials of degree at most 4 with coefficients in the field $\mathbf{F}$. The mapping

$$\mathbf{T}(a_0 + a_1 x + \cdots + a_4 x^4) = (a_0, a_1, \ldots, a_4)$$

is an isomorphism between $\mathbb{P}_4[x]$ and $\mathbf{F}^5$.

Similarly, it is easy to define isomorphisms between $\mathbf{M}_n(\mathbf{F})$ and $\mathbf{F}^{n^2}$.

Examples

- A linear transformation $\mathbf{T} : \mathbf{V} \to \mathbf{V}$ such that $\mathbf{T}^2 = 0$ obviously cannot be invertible. Note however that $\mathbf{I} - \mathbf{T}$ is always invertible:

$$(\mathbf{I} - \mathbf{T})(\mathbf{I} + \mathbf{T}) = \mathbf{I} - \mathbf{T}^2 = \mathbf{I}.$$

- Prove the same assertion if $\mathbf{T}^3 = 0$ [or any other power $\mathbf{T}^n = 0$].
- Let $\mathbf{V}$ be the vector space of all sequences $(a_1, a_2, a_3, \ldots)$. The functions right shift and left shift are L.T.

$$\begin{aligned} \mathbf{r}(a_1, a_2, a_3, \ldots) &= (0, a_1, a_2, \ldots) \\ \mathbf{s}(a_1, a_2, a_3, \ldots) &= (a_2, a_3, a_4 \ldots) \end{aligned}$$

$\mathbf{r}$ is one-one but not an isomorphism, $\mathbf{s}$ is onto but not an isomorphism

Exercise 5: Let $\mathbf{A}$ be a $n \times n$ of rank r. Define the mapping $\mathbf{T} : \mathbf{M}_n(\mathbf{F}) \to \mathbf{M}_n(\mathbf{F})$ by

$$\mathbf{B} \mapsto \mathbf{AB}.$$

Show that $\mathbf{T}$ is a linear transformation of rank $r \cdot n$.

Exercise 6: Show that there is no square nonzero real matrix $\mathbf{A}$ such that

$$\mathbf{A}^t = r\mathbf{A}, \quad r \neq \pm 1.$$

Change of coordinates

Quickly: Changing coordinates permit the solution of many problems. Here are two:

- To evaluate $\int_0^1 te^{t^2}dt$, one sets $y = t^2$ and the problem becomes

$$\int_0^1 te^{t^2}dt = \int_0^1 \frac{1}{2}e^y dy = \frac{1}{2}(e-1).$$

- What is the graph of the equation $2x^2 + 6xy + 10y^2 = 100$? The solution requires a change of point-of-view–which a change of coordinates will bring.

The change of coordinates issue we will discuss is the following: Let $v \in \mathbf{V}$ be a vector of a V.S. If two bases $\mathcal{A} = v_1, \ldots, v_n$ and $\mathcal{B} = u_1, \ldots, u_n$ are picked in $\mathbf{V}$, the vector has two representations

$$[v]_{\mathcal{A}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad [v]_{\mathcal{B}} = \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}$$

Question: How are the x_i related to the x_i'? The answer will depend on how the v_i and u_i relate.

Change of bases formula

We have

$$v = \sum_{j=1}^{n} x_j v_j = \sum_{j=1}^{n} x'_j u_j.$$

We start from

$$v_j = \sum_{i=1}^{n} p_{ij} u_i, \quad u_j = \sum_{i=1}^{n} q_{ij} v_i$$

Note the two [basis changing] matrices

$$\mathbf{P} = [p_{ij}], \quad \mathbf{Q} = [q_{ij}]$$

If we replace $v_j = \sum_{i=1}^{n} p_{ij} u_i$ in

$$v = \sum_{j=1}^{n} x_j v_j = \sum_{j=1}^{n} x'_j u_j.$$

we get

$$v = \sum_{j=1}^{n} x_j (\sum_{i=1}^{n} p_{ij} u_i) = \sum_{i=1}^{n} (\sum_{j=1}^{n} p_{ij} x_j) u_i = \sum_{i=1}^{n} x'_i u_i.$$

$$x_i = \sum_{j=1}^{n} p_{ij} x'_j,$$

the desired formulas.

In matrix notation:

$$\begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} = \mathbf{Q} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{P} \cdot \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}$$

Note

$$\mathbf{P} \cdot \mathbf{Q} = \mathbf{I}$$

Max and Min of functions of several variables

The need for change of variables occur in the determination of local maxima and minima of functions of several variables. Recall that the function $\mathbf{f}(x, y)$ has a local maximum at (a, b) if

$$\mathbf{f}(a, b) \geq \mathbf{f}(x, y),$$

for (x, y) near (a, b). If $\mathbf{f}$ has derivatives, this first requires

$$\frac{\partial \mathbf{f}}{\partial x}(a, b) = \frac{\partial \mathbf{f}}{\partial y}(a, b) = 0$$

What else? We expand $\mathbf{f}$ around (a, b):

$$\begin{aligned} \mathbf{f}(x, y) &= \mathbf{f}(a, b) + (\underbrace{\mathbf{f}_x(a, b)(x - a) + \mathbf{f}_y(a, b)(y - b) = 0}) \\ &+ \underline{1/2(\mathbf{f}_{xx}(a, b)(x - a)^2 + 2\mathbf{f}_{xy}(x - a)(y - b) + \mathbf{f}_{yy}(y - b)^2)} + \end{aligned}$$

Whether (a, b) is a local maximum will depend on whether the term

$$\underline{1/2(\mathbf{f}_{xx}(a,b)(x-a)^2 + 2\mathbf{f}_{xy}(x-a)(y-b) + \mathbf{f}_{yy}(y-b)^2)}$$

is always non-positive near (a, b).
Hard to guess when a polynomial

$$Ax^2 + Bxy + Cy^2$$

is always negative near the origin, UNLESS $B = 0$ when the condition is $A, C \leq 0$. It involves the examination of

$$\begin{bmatrix} \mathbf{f}_{xx} & \mathbf{f}_{xy} \\ \mathbf{f}_{yx} & \mathbf{f}_{yy} \end{bmatrix}$$

Just imagine the size of the problem in 5 or 10 variables! Fortunately, Linear Algebra comes to the rescue: it involves a certain calculation with the matrix of second order derivatives. In the case of 3 variables,

$$\begin{bmatrix} \mathbf{f}_{xx} & \mathbf{f}_{xy} & \mathbf{f}_{xz} \\ \mathbf{f}_{yx} & \mathbf{f}_{yy} & \mathbf{f}_{yz} \\ \mathbf{f}_{zx} & \mathbf{f}_{zy} & \mathbf{f}_{zz} \end{bmatrix}$$

This is so important that we will have to return to the topic for a serious treatment.

Outline

Rings

A ring R is a set with two composition laws, called 'addition' and 'multiplication', say $+$ and $\times$: $\forall a, b \in R$ have compositions $a + b$ and $a \times b$. (The second composition is also written $a \cdot b$, or simply ab.)

- $(R, +)$ is an abelian group
- $(R, \times)$: multiplication is associative, and distributive over $+$, that is $\forall a, b, c \in R$,

$$(ab)c = a(bc), \quad ab = ba, \quad a(b + c) = ab + ac$$

- existence of identity: $\exists e \in R$ such that

$$\forall a \in R \quad e \times a = a \times e = a$$

- If $ab = ba$ for all $a, b \in R$, the ring is called commutative

There is a unique identity element e, usually we denote it by 1:

$$e = ee' = e'e = e'$$

Field

A field $\mathbf{F}$ is a set with two composition laws, called 'addition' and 'multiplication', say $+$ and $\times$: $\forall a, b \in \mathbf{F}$ have compositions $a + b$ and $a \times b$. (The second composition is also written $a \cdot b$, or simply ab.)

- $(\mathbf{F}, +)$ is an abelian group
- $(\mathbf{F}, \times)$: multiplication is associative, commutative and distributive over $+$, that is $\forall a, b, c \in \mathbf{F}$,

$$(ab)c = a(bc), \quad ab = ba, \quad a(b + c) = ab + ac$$

- existence of identity $\exists e \in \mathbf{F}$ such that

$$\forall a \in \mathbf{F} \quad a \times e = a$$

- existence of inverses For every $a \neq 0$, there is $b \in \mathbf{F}$

$$a \times b = e.$$

There is a unique element e, usually we denote it by 1. For $a \neq 0$, the element b such that $ab = 1$ is unique; it is often denoted by $1/a$ or a^{-1}.

We can now define scalars: the elements of a field.

Fields are ubiquotous:

- $\mathbb{R}$: real numbers
- The integers $\mathbb{Z}$ is not a field (not all integers have inverses), but $\mathbb{Q}$, the rational numbers is a field.
- $\mathbb{C}$: complex numbers, $z = a + bi$, $i = \sqrt{-1}$, with compositions

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

$$(a + bi) \times (c + di) = (ac - bd) + (ad + bc)i$$

The arithmetic here requires a bit more care:

If $a + bi \neq 0$,

$$\frac{1}{a+bi} = \frac{a-bi}{a^2+b^2} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2} i$$

Exercise: Number fields

Let **F** be the set of all real numbers of the form

$$z = a + b\sqrt{2}, \quad a, b \in \mathbb{Q}$$

prove that **F** is a field.

Query: How to prove a subset **F** of the field $\mathbb{R}$ is a field?

Another noteworthy example is $\mathbb{F}_2$, the set made up by two elements $\{0, 1\}$ (or (even, odd))with addition defined by the table

$+$	0	1
0	0	1
1	1	0

$$1 + 1 = 0!$$

and multiplication by

$\times$	0	1
0	0	0
1	0	1

Exercise 1: Prove that in any field **F** the rule minus times minus is plus holds, that is for any $a, b \in \mathbf{F}$,

$$-(-a) = a, \quad (-a)(-b) = ab.$$

Solution: The first assertion folllows from

$a + (-a) = (-a) + a = O$.

Because of the above, we must show that $(-a)(-b)$ is the negative of $-(ab)$. We first claim $(-a)b = -(ab)$. Note

$$(-a)b + ab = ((-a) + a)b = Ob = O.$$

$$(-a)(-b) - (ab) = (-a)(-b) + (-a)b = (-a)((-b) + b) = (-a)O = O.$$

A field is the mathematical structure of choice to do arithmetic. Given a field **F**, fractions can defined as follows: If $a, b \in \mathbf{F}, \quad b \neq 0,$

$$\frac{a}{b} := ab^{-1}.$$

The usual calculus of fractions then follows, for instance

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

Homework #1

1. Let **G** be a group such that $x^2 = 1$ for all $x \in \mathbf{G}$. Prove that **G** is abelian.
2. Prove that in any group the orders of ab and of ba are equal.
3. Prove that the set of elements of finite order in an abelian group is a subgroup.
4. An nth root of unity is a complex z such that $z^n = 1$. Prove that the nth roots of unity form a cyclic subgroup of $\mathbb{C}^\times$ of order n. If $n = 12$, determine the number of generators of this subgroup.

Exercise

Let **G** be the group of 3×3 invertible matrices with entries in $\mathbb{Z}_2$. What is the order of **G**?
Every $\mathbf{A} \in \mathbf{G}$ is a matrix

$$\mathbf{A} = [v_1 | v_2 | v_3]$$

whose column vectors are linearly independent. Let us count:

- v_1 can be any nonzero vector of $\mathbb{Z}_2^3$: $2^3 - 1$ choices
- v_2 can be any vector which is not a multiple of v_1: $2^3 - 2$ choices
- v_3 can be any vector which is not a linear combination of v_1, v_2: $2^3 - 2^2$ choices

$$|\mathbf{G}| = 7 \cdot 6 \cdot 4$$

Word processing

Let **G** be a group and $\mathbb{X} = \{a, b, c, \ldots\}$ a subset. Using $\mathbb{X}$ as an alphabet, consider the set of all products

$$a_1 \cdot a_2 \cdots a_n$$

where $a_i \in \mathbb{X}$ or $a_i^{-1} \in \mathbb{X}$.
This set, $\langle \mathbb{X} \rangle$, is group, the group generated by $\mathbb{X}$.
Word prcessing in groups is ...

Homework #1 extra

- Describe all groups of order 4.
- How many groups of order 8 can you name?
- How many groups of order 6 can you name?

Outline

Relations

Definition

Let A and B be sets. R is a **relation from** A **to** B iff R is a subset of $A \times B$,

$$R \subset A \times B.$$

If $(a, b) \in R$ we write $a\ R\ b$ and say that a is R-related to b. If $(a, b) \notin R$, we write $a \not R\ b$. A relation R from A to A is called a **relation of** A: $R \subset A \times A$.

There are many notations for relations: familiar ones are $a \simeq b$, $a \geq b$, $a \mid b$, etc.

Example: I_A (**identity**) of A is the relation $a \simeq b$ iff $a = b$.

Another: $a\ R\ b$ for all $a, b \in A$.

Equivalence Relations

Definition

Let A be a set and R a relation on A.

- R is reflexive iff for all $x \in A$, $x \; R \; x$.
- R is symmetric iff for all $x \in A$ and $y \in A$, if $x \; R \; y$, then $y \; R \; x$.
- R is transitive iff for all x, y and z in A, if $x \; R \; y$ and $y \; R \; z$, then $x \; R \; z$.

Example

Let R be the set of all $(x, y) \in \mathbb{N} \times \mathbb{N}$ such that $x + y$ is divisible by 3. Is this relation symmetric? reflexive? transitive?

1. If $(x, y) \in R$, $x + y = 3m$, for some m. Then $y + x = 3m$ also, so $(y, x) \in R$: **Symmetric**
2. $(1, 1) \notin R$: **Not Reflexive**
3. $(1, 2), (2, 1) \in R$ but $(1, 1) \notin R$: **Not Transitive**

Equivalence Relation

Definition

A relation R on a set A is an **equivalence relation on** A iff R is reflexive, symmetric, and transitive.

Let R be a relation on the set A.

- **R Reflexive**: $a \to a \, \forall a \in A$
- **R Symmetric**: $a \to b \Rightarrow b \to a$
- **R Transitive**: $a \to b$ and $b \to c \Rightarrow a \to c$

Equivalence Class

Definition

Let R be an equivalence relation on the set A. For $x \in A$, the **equivalence class of** x determined by R is the set

$$x/R = \{y \in A : x \; R \; y\}.$$

This is read "the class of x modulo R." The set of all equivalence classes of R is called A **modulo** R and denoted $A/R = \{x/R : x \in A\}$.

Example: Two integers have the same **parity** if they are both even or both odd. Let $R = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} : x \text{ and } y \text{ have the same parity.}\}$ R is an equivalence relation with two equivalence classes: the even integers E and the odd integers D. $\mathbb{Z}/R = \{E, D\}$.

Big Example

Let m be a fixed, nonzero integer. Let $\equiv_m$ be the relation on $\mathbb{Z}$,

$$x \equiv_m y \text{ iff } m \text{ divides } x - y.$$

This is also written $x \equiv y \pmod{m}$ or even $x = y \pmod{m}$. It is easy to see that $\mathbb{Z}/\equiv_2 = \{E, D\}$. This set is also denoted by $\mathbb{Z}_2$ and called the set of integers modulo 2. For $m = 3$, $\equiv_3$ is also an equivalence relation and there are three distinct equivalence classes.

Theorem

The relation $\equiv_m$ is an equivalence relation on the integers. The set of equivalence relations is called $\mathbb{Z}_m$ and has m distinct elements $\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{m-1}$.

Partitions

Definition

Let A be a nonempty set. A **partition** of A is a set $\mathcal{A}$ of subsets of A such that

1. If $X \in \mathcal{A}$, then $X \neq \emptyset$.
2. If $X \in \mathcal{A}$ and $Y \in \mathcal{A}$, $X \neq Y$, then $X \cap Y = \emptyset$.
3. $\bigcup_{X \in \mathcal{A}} X = A$.

How do partitions arise? It will be a pretty straight answer.

Example

Let $A = \{a, b, c\}$–the following are partitions of A:

- $\{\{a, b, c\}\}$
- $\{\{a\}, \{b\}, \{c\}\}$,
- $\{\{a\}, \{b, c\}\}$,
- $\{\{b\}, \{a, c\}\}$
- $\{\{c\}, \{a, b\}\}$

Are they all?

Partitions versus Equivalence Classes

Theorem

Let $\mathcal{B}$ be a partition of the nonempty set A. For x and y in A, define $x\ Q\ y$ iff there exists $C \in \mathcal{B}$ such that $x \in C$ and $y \in C$. Then

1. *Q is an equivalence relation on A.*
2. *$A/Q = \mathcal{B}$.*

Example: Define the following sets of $\mathbb{Z}$:

$$\begin{aligned} A_0 &= \{3k : k \in \mathbb{Z}\} = \{\ldots, -6, -3, 0, 3, 6, \ldots\} \\ A_1 &= \{3k+1 : k \in \mathbb{Z}\} = \{\ldots, -5, -2, 1, 4, 7, \ldots\} \\ A_2 &= \{3k+2 : k \in \mathbb{Z}\} = \{\ldots, -4, -1, 2, 5, 8, \ldots\} \end{aligned}$$

The partition defines the relation we denoted $\equiv_3$.

How Partitions arise

Theorem

Let R be an equivalence relation on a nonempty set A. Then

1. *For all $x \in A$, $x/R \subseteq A$ and $x \in x/R$. (Thus $x/R \neq \emptyset$.*
2. $\bigcup_{x \in A} x/R = A$.
3. $x \; R \; y$ *iff* $x/R = y/R$.
4. $x \not R \; y$ *iff* $x/R \cap y/R = \emptyset$.

Thus, the set $\{x/R : x \in A\}$ of equivalence classes is a partition of A.

In words: An equivalence relation on the set A gives rise to a partition of A and vice-versa.

Outline

Cosets

Definition

Let **H** be a subgroup of the group **G**. A left coset of **H** is a subset of the form

$$a\mathbf{H} = \{ah : h \in \mathbf{H}\}.$$

Proposition

The left cosets of **H** *are the equivalence classes for the congruence relation:*

$$a \equiv b : \text{if } b = ah, \text{ for some } h \in \mathbf{H}.$$

Proof. Volunteers please!

Corollary

The left cosets of a subgroup partition the group.

Index of a subgroup

Among the properties of this construction:

- The correspondence

$$\mathbf{H} \to a\mathbf{H} : h \to ah$$

  is $1-1$ onto, so the cardinality of $\mathbf{H}$ and of the coset $a\mathbf{H}$ are equal.

- The partition of $\mathbf{G}$ defined by the coset decomposition

$$\mathbf{G} = \bigcup a\mathbf{H}$$

  gives rise to important numerical relationships.

- The number of cosets is denoted $[\mathbf{G} : \mathbf{H}]$ and called the index of the subgroup.

Lagrange Theorem

Theorem (Counting Formula)

If $\mathbf{H}$ *is a subgroup of the finite group* $\mathbf{G}$*,*

$$|\mathbf{G}| = |\mathbf{H}| \cdot [\mathbf{G} : \mathbf{H}].$$

Theorem (Lagrange Theorem)

Let $\mathbf{G}$ *be a finite group, and let* $\mathbf{H}$ *be a subgroup of* $\mathbf{G}$*. Then the order of* $\mathbf{H}$ *divides the order of* $\mathbf{G}$*.*

Corollary

If $\mathbf{G}$ *is a group of order* p*, a prime, then any* $a \neq 1$ *has order* p*. In particular, all groups of order* p *are isomorphic.*

Proof. If $a \neq 1$, the order of the subgroup $\langle a \rangle$ is p.

Example

Lagrange formula is helpful when we seek the list the subgroups of a group.
Let $\mathbf{G} = S_3$, the symmetric group of three letters. Since $\mathbf{G}$ has order 6, the intermediate subgroups have order 2 or 3.

1. Order 1: (1)
2. Order 6: S_3
3. Order 2: 3 subgroups: $\langle(1,2)\rangle$, $\langle(1,3)\rangle$, $\langle(2,3)\rangle$
4. Order 3: The subgroup generated by the permutation: $1 \to 2 \to 3 \to 1$

Example

In the case S_4, its order, 24, has many divisors. With more technology, we will be able do describe them. This is a promise!

Example

Dimension Formula

Recall:

Theorem

Let $\mathbf{T} : \mathbf{V} \to \mathbf{W}$ *be a linear transformation of finite dimensional vector spaces. Then*

$$\dim N(\mathbf{T}) + \dim R(\mathbf{T}) = \dim \mathbf{V}.$$

That is, nullity $+$ *rank* $= \dim \mathbf{V}$.

A group theoretic version:

Theorem

Let $\varphi : \mathbf{G} \to \mathbf{G}'$ *be a homomorphism of finite groups. Then*

$$|\mathbf{G}| = |\ker \varphi| \cdot |\mathrm{im}\, \varphi|.$$

Proof. Just observe the $1 - 1$ onto correspondence

Right cosets

Definition

Let **H** be a subgroup of the group **G**. A right coset of **H** is a subset of the form

$$\mathbf{H}a = \{ha : h \in \mathbf{H}\}.$$

Proposition

The right cosets of **H** *are the equivalence classes for the congruence relation:*

$$a \equiv b : \text{if } b = ha, \text{ for some } h \in \mathbf{H}.$$

Normal subgroups

Proposition

A subgroup $\mathbf{H}$ *of a group* $\mathbf{G}$ *is normal if and only if every left coset is also a right coset. If* $\mathbf{H}$ *is normal, then* $a\mathbf{H} = \mathbf{H}a$ *for every* $a \in \mathbf{G}$.

Proof. Volunteer!

Example

Definition

Let **G** be a group. The commutator of $a, b \in \mathbf{G}$ is the element

$$aba^{-1}b^{-1}.$$

It is denoted $[a, b]$.

- $[a, b]^{-1} = [b, a]$
- **Achtung:** The product of two commutators may not be a commutator.
- If $\phi : \mathbf{G} \to \mathbf{H}$ is a group homomorphism and **H** is an abelian group, then all $[a, b] \in \ker \phi$.

Definition/Exercise

- Let $\mathbf{G}$ be a group, and $\mathbf{G}^{(1)}$ the set of all finite products of commutators,

$$[a_1, b_1][a_2, b_2] \cdots [a_n, b_n].$$

  Prove that $\mathbf{G}^{(1)}$ is a normal subgroup of $\mathbf{G}$. $\mathbf{G}^{(1)}$ is called the commutator subgroup of $\mathbf{G}$.
- If $\mathbf{G} = S_3$, determine $\mathbf{G}^{(1)}$.

Outline

Homework #2

1. Find all the subgroups of S_3, and determine which are normal. [If you need a challenge, do same for S_4.]
2. Find all the subgroups of the quaternion group, and determine which are normal.
3. Prove that the center of a group is a normal subgroup.
4. Let $\varphi, \phi : \mathbf{G} \to \mathbf{G}'$ be two group homomorphisms, and let $H \subset \mathbf{G}$ be the subset

$$\{x \in \mathbf{G} : \varphi(x) = \phi(x)\}$$

   Prove or disprove: **H** is a subgroup.
5. If **H** is a subgroup of **G** of index 2, prove that **H** is normal.

Outline

Products of Groups

Definition

Let $\mathbf{G}$ and $\mathbf{G}'$ be two groups and $\mathbf{G} \times \mathbf{G}'$ the product set. The composition

$$(a, a'), (b, b') \mapsto (ab, a'b')$$

is a group law on $\mathbf{G} \times \mathbf{G}'$ called the product group of $\mathbf{G}$ and $\mathbf{G}'$.

Example

Let $\mathbf{G}$ and $\mathbf{G}'$ be two copies of $\mathbb{Z}_2$. The group $\mathbf{V}$ that Jonathan discussed in last class was obtained as

$$\mathbf{V} = \{(a, b) : a, b \in \mathbb{Z}_2\} = \mathbf{G} \times \mathbf{G}'.$$

Properties

It is defined by the inclusion and projection homomorphisms

$$\begin{tikzcd}
\mathbf{G} \arrow[dr, "i"] & & \mathbf{G} \\
& \mathbf{G} \times \mathbf{G}' \arrow[ur, "p"] \arrow[dr, "p'"] & \\
\mathbf{G}' \arrow[ur, "i'"] & & \mathbf{G}'
\end{tikzcd}$$

- $i(x) = (x, 1')$
- $i'(x') = (1, x')$
- $p(x, x') = x$
- $p'(x, x') = x'$
- $(1, 1')$ is the identity of $\mathbf{G} \times \mathbf{G}'$
- $(a, b)^{-1} = (a^{-1}, b^{-1})$

Mapping Property of Products

Proposition

Let $\mathbf{H}$ *be any group. The homomorphisms* $\Phi : \mathbf{H} \to \mathbf{G} \times \mathbf{G}'$ *are in bijective correspondence with pairs* (ϕ, ϕ') *of homomorphisms*

$$\phi : \mathbf{H} \to \mathbf{G}, \quad \phi' : \mathbf{H} \to \mathbf{G}'.$$

The kernel of Φ *is* $\ker \phi \cap \ker \phi'$.

Proof. Given a pair (ϕ, ϕ') of homomorpisms, the mapping

$$\Phi(a) = (\phi(a), \phi'(a)) \in \mathbf{G} \times \mathbf{G}', \quad a \in \mathbf{H}$$

is clearly a homomorphism.

Conversely, given Φ, define the homomorphisms $\phi = p \circ \Phi$ and $\phi' = p' \circ \Phi$.
Finally, note that $\Phi(h) = (1, 1')$ if and only if $\phi(h) = 1$ and $\phi'(h) = 1'$. Thus the kernel of Φ is $\ker \phi \cap \ker \phi'$.

Example

If C_6 is a cyclic group of order 6, $C_6 = \{1, x, x^2, x^3, x^4, x^5\}$, and $\mathbf{G} = \{1, x^3\}$ and $\mathbf{G}' = \{1, x^2, x^4\}$ then

$$C_6 \simeq C_2 \times C_3$$

Discussion:

- $C_2 \times C_3$ is a group of order 6.
- The element $z = (x^3, x^2)$ has order 6: The powers $1, z, z^2, z^3, z^4, z^5$ are distinct.
- Thus $C_2 \times C_3$ is cyclic so it isomorphic to C_6.

General example...

Proposition

Let r, s be integers with no common (± 1) factor. Then a cyclic group of order rs is isomorphic to the direct product of a cyclic group of order r and a cyclic group of order s.

Discussion: Reminds you of another proof? Note:

- $C_{rs} = \langle x \rangle$, $C_r = \langle x^s \rangle$, $C_s = \langle x^r \rangle$
- cconsider $z = (x^s, x^r)$
- Find its order, that is the smallest integer n such that $z^n = (1, 1)$. Do it!

Product Groups

Definition

Let A and B be subsets of a group $\mathbf{G}$. Then

$$AB = \{ab : a \in A, b \in B\}.$$

Proposition

Let $\mathbf{H}$ *and* $\mathbf{K}$ *be subgroups of a group* $\mathbf{G}$.

1. *If* $\mathbf{H} \cap \mathbf{K} = \{1\}$, *the product* $p : \mathbf{H} \times \mathbf{K} \to \mathbf{G}$ *defined by* $p(h, k) = hk$ *is injective. Its image is the subset* $\mathbf{HK}$.
2. *If either* $\mathbf{H}$ *or* $\mathbf{K}$ *is a normal subgroup of* $\mathbf{G}$, *then* $\mathbf{HK} = \mathbf{KH}$ *and* $\mathbf{HK}$ *is a subgroup of* $\mathbf{G}$.
3. *If* $\mathbf{H}$ *and* $\mathbf{K}$ *are normal subgroups,* $\mathbf{H} \cap \mathbf{K} = \{1\}$, *and* $\mathbf{HK} = \mathbf{G}$, *then* $\mathbf{G}$ *is isomorphic to the product.*

Proof

- Let (a_1, b_1) and (a_2, b_2) elements of $\mathbf{H} \times \mathbf{K}$ such that $a_1 b_1 = a_2 b_2$. Rewrite the equation as $a_2^{-1} a_1 = b_2 b_1^{-1}$. Since $\mathbf{H} \cap \mathbf{K} = \{1\}$, $a_2^{-1} a_1 = b_2 b_1^{-1} = 1$, hence $a_1 = a_2$ and $b_1 = b_2$. This shows that p is injective.
- Suppose $\mathbf{H}$ is a normal subgroup of $\mathbf{G}$. Suppose that $a \in \mathbf{H}$ and $b \in \mathbf{K}$. Then $bab^{-1} \in \mathbf{H}$, so

$$ba = bab^{-1} b \in \mathbf{HK}.$$

This shows that $\mathbf{KH} \subset \mathbf{HK}$. The reverse inclusion has similar proof. That $\mathbf{HK}$ is a group follows quickly: do in class.

Proof cont'd

- Suppose both **H** and **K** are normal subgroups of **G** and $\mathbf{H} \cap \mathbf{K} = \{1\}$. If $a \in \mathbf{H}$ and $b \in \mathbf{K}$, consider

$$(aba^{-1})b^{-1} = a(ba^{-1}b^{-1}).$$

  The first term lies in **H**, the second in **K**. This proves that $ab = ba$.
  It follows by direct verification that the product mapping $p : \mathbf{H} \times \mathbf{K} \to \mathbf{G}$ is an injective homomorphism.
  The assumption that $\mathbf{HK} = \mathbf{G}$ implies that $\mathbf{H} \times \mathbf{K} \simeq \mathbf{G}$.

Outline

Modular Arithmetic

Let n be a fixed positive integer. Let us describe an invention of Gauss. Recall:

Definition

Two integers, a, b are said to be congruent modulo n,

$$a \equiv b, \quad \text{modulo } n,$$

if n divides $b - a$.

This is an equivalence relation, and the equivalence classes are called congruence classes modulo n, or residue classes modulo n.

The congruence class of the integer a is denoted $\overline{a}$ (or $[a]$)

$$\overline{a} = \{\ldots, a-2n, a-n, a, a+n, a+2n, \ldots\}.$$

The congruence class of 0 is the subgroup $n\mathbb{Z}$,

$$\overline{0} = \{\ldots, -2n, -n, 0, n, 2n, \ldots\}.$$

Note that the congruence classes modulo n are the cosets of the subgroup $n\mathbb{Z}$ of $\mathbb{Z}$.

Proposition

There are n congruence classes modulo n,

$$\overline{0}, \overline{1}, \ldots, \overline{n-1}.$$

Or, the index $[\mathbb{Z} : n\mathbb{Z}]$ *of the subgroup* $n\mathbb{Z}$ *in* $\mathbb{Z}$ *is n.*

Carl Friedrich Gauss (1777-1855)

New Composition Laws

Pseudo definition!

Definition

Let $\overline{a}$ and $\overline{b}$ be two congruence classes modulo n.

$$\begin{aligned} \overline{a} + \overline{b} &:= \overline{a+b} \\ \overline{a}\overline{b} &:= \overline{ab} \end{aligned}$$

To justify:

Lemma

If $a \equiv a'$ and $b \equiv b'$ (modulo n), then $a + b \equiv a' + b'$ (modulo n) and $ab \equiv a'b'$ (modulo n).

Proof.

Assume $a' \equiv a$ and $b' \equiv b$ (modulo n). Then $a' = a + rn$ and $b' = b + sn$ for some integers r, s. Then

$$\begin{aligned} a' + b' &= (a + rn) + (b + sn) = a + b + (r + s)n, \\ a'b' &= (a + rn)(b + sn) = ab + (rb + sa + sn)n. \end{aligned}$$

□

The associative, commutative and distributive laws will hold, that makes the congruence classes modulo n a ring: $\mathbb{Z}/n\mathbb{Z}$, or $\mathbb{Z}_2$.

Example

Perhaps the noteworthy example is $\mathbb{F}_2$, the set made up by two elements $\{0, 1\}$ (or (even, odd))with addition given by the table

$+$	0	1
0	0	1
1	1	0

$$1 + 1 = 0!$$

and multiplication by

$\times$	0	1
0	0	0
1	0	1

$\mathbb{Z}/(p)$

Let p be a prime number. The ring $\mathbb{Z}/(p)$ is a group for the addition operation and the subset **F** of cosets

$$\{1, 2, \ldots, p-1\}$$

is a group for the multiplication operation.

Theorem (Wilson's Theorem)

The integer p is a prime if and only if

$$(p-1)! = -1 \mod p.$$

Proof. Note that $(p-1)!$ mod p is the product of all elements of the group **F**. If $x \neq -1$ in **F**, both x and its inverse occur as factors. Thus $(p-1)! = -1$ in **F**.

- If p is composite, $p = ab$ and $a < b$, then ab divides $(p-1)!$.
- If $a = b$, then $p = a^2$. If $a = 2$, then

$$(a^2 - 1)! = 6 \equiv 2 \bmod 4$$

  but $2 \not\equiv -1 \bmod 4$.
- If $2 < a$, then $2a < a^2$ and so a and $2a$ are factors of $(a^2 - 1)!$; therefore $(a^2 - 1)! \equiv 0 \bmod a^2$, and the proof is complete.

Outline

Homework #3

1. Let **G** be a group containing normal subgroups of order 3 and 5. Prove that **G** contains a normal subgroup of order 15.
2. Determine the integers n for which the congruences $x + y \equiv 2$, $2x - 3y \equiv 3$ (modulo n) have a solution.
3. Let $\mathbf{H} = \{\pm 1, \pm i\}$ be the subgroup of $\mathbf{G} = \mathbb{C}^*$ of fourth roots of the unity. Describe the cosets of **H** in **G** explicitly, and prove that $\mathbf{G}/\mathbf{H}$ is isomorphic to **G**. Generalize (means what?).
4. Prove that the groups $\mathbb{R}^+/\mathbb{Z}^+$ and $\mathbb{R}^+/2\pi\mathbb{Z}^+$ are isomorphic.

Outline

Last Class ... and ... Today

- Product Groups: Last Time
- Quotient Groups: Today

Outline

Quotient Groups

Let us begin with the discussion of a nice example.

Let $\mathbb{R}^2$ be the usual real plane and let **L** be a line passing through the origin. [Carry an example in your mind.] **L** is a subspace of $\mathbb{R}^2$.

For any vector $v \in \mathbb{R}^2$, $v + \mathbf{L}$ is the set obtained by translating **L** by v. It is a line parallel to **L**. We are going to denote it $\mathbf{L}_v$ and the set of all such such lines we denote by **V**= all lines parallel to **L**.

A feature of the notation $\mathbf{L}_v$ is the following. Suppose $u \in \mathbf{L}$. Then $\mathbf{L}_u = \mathbf{L}_O = \mathbf{L}$. More generally,

$$\mathbf{L}_v = \mathbf{L}_{v+u}$$

v is said to be a representative of $\mathbf{L}_v$, but the observation says that $v + u$ is also a representative of $\mathbf{L}_v$. Essentially any vector in $\mathbf{L}_v$ serves as its representative.

This will be cause for confusion!

Let us define an 'addition' for this set of lines. We are going to use the ordinary '$+$' when '$\oplus$' or '$\otimes$' might be more cautious. For any two lines $\mathbf{L}_v$ and $\mathbf{L}_w$

$$\mathbf{L}_v + \mathbf{L}_w := \mathbf{L}_{v+w}.$$

For instance $\mathbf{L}_v + \mathbf{L}_O = \mathbf{L}_v$.

Proposition

This composition does not depend of the representatives used, that is if $\mathbf{L}_v = \mathbf{L}_{v'}$ *and* $\mathbf{L}_w = \mathbf{L}_{w'}$ *then*

$$\mathbf{L}_v + \mathbf{L}_w = \mathbf{L}_{v'} + \mathbf{L}_{w'}.$$

It is obvious now that this composition is commutative and associative. The line $\mathbf{L}$ plays the role of O: $\mathbf{L}_v + \mathbf{L} = \mathbf{L}_v$, and $\mathbf{L}_{-v}$ is the negative of $\mathbf{L}_v$: $\mathbf{L}_v + \mathbf{L}_{-v} = \mathbf{L}$.

If we define scalar multiplication by

$$c\mathbf{L}_v = \mathbf{L}_{cv}$$

it will check easily that $\mathbf{V}$ is a vector space. It is called the quotient of $\mathbb{R}^2$ by $\mathbf{L}$. The notation $\mathbf{V} = \mathbb{R}^2/\mathbf{L}$ is used.

Quotient Groups

The following observation gives a composition law on the cosets of a normal subgroup of a group.

Lemma

Let $\mathbf{N}$ *be a normal subgroup of a group* $\mathbf{G}$*. Then the product of two cosets* $a\mathbf{N}$ *and* $b\mathbf{N}$ *is a coset, in fact*

$$(a\mathbf{N})(b\mathbf{N}) = ab\mathbf{N}.$$

Proof. Recall that $a\mathbf{N} = \mathbf{N}a$. Then

$$(a\mathbf{N})(b\mathbf{N}) = a(\mathbf{N}b)\mathbf{N} = a(b\mathbf{N})\mathbf{N} = ab(\mathbf{N}\mathbf{N}) = ab\mathbf{N}.$$

- We denote the set of cosets by $\mathbf{G}/\mathbf{N}$.
- The coset $(\mathbf{N})$ acts as the identity: $(a\mathbf{N})(\mathbf{N}) = (\mathbf{N})(a\mathbf{N}) = (a\mathbf{N})$.
- The inverse of $(a\mathbf{N})$ is $(a^{-1}\mathbf{N})$.

Quotient Group

The assignment $a \in \mathbf{G} \to a\mathbf{N} \in \mathbf{G}/\mathbf{H}$ is denoted $\pi : \mathbf{G} \to \mathbf{G}/\mathbf{N}$. These observations are summarized in:

Theorem

$\overline{\mathbf{G}} = \mathbf{G}/\mathbf{N}$ is a group and $\pi : \mathbf{G} \to \mathbf{G}/\mathbf{N}$ is a group homomorphsm whose kernel is $\mathbf{N}$. The order of $\overline{\mathbf{G}}$ is the index $[\mathbf{G} : \mathbf{N}]$ of $\mathbf{N}$ in $\mathbf{G}$.

Proof.

- Verify that $\overline{\mathbf{G}}$ is a group.
- Verify that π is a group homomorphism of kernel $\mathbf{N}$.

First Isomorphism Theorem

Theorem (First Isomorphism Theorem)

Let $\varphi : \mathbf{G} \to \mathbf{G}'$ *be a surjective homomorphism, and let* $\mathbf{N} = \ker \varphi$. *Then the mapping*

$$\overline{\varphi} : \mathbf{G}/\mathbf{N} \to \mathbf{G}'$$

defined by $\overline{\varphi}(a\mathbf{N}) = \varphi(a)$ *is group isomorphism.*

Proof. Main points:

- Verify $\overline{\varphi}$ is well defined. If $a\mathbf{N} = b\mathbf{N}$, then $\varphi(a) = \varphi(b)$: For any $c \in \mathbf{N}$, $\varphi(ac) = \varphi(a)$ because $\varphi(c) = 1'$.
- Verify that $\overline{\varphi}$ is an injective homomorphism: If $\overline{\varphi}(a\mathbf{N}) = 1'$, $a \in \mathbf{N}$.

Finitely Generated Abelian Groups

Definition

A group **G** is finitely generated if there is a finite set of elements in **G**, $\{a_1, \ldots, a_n\}$, such that every element $z \in \mathbf{G}$ is a product of the a_i and/or of their inverses. $\{a_1, \ldots, a_n\}$ is called a set of generators of **G**. (A cyclic group $\langle x \rangle$ is an example.)

- Notation: $\mathbf{G} = \langle a_1, \ldots, a_n \rangle$.
- Major example is the group

$$\mathbb{Z}^n = \underbrace{\mathbb{Z} \times \cdots \times \mathbb{Z}}_{n \text{ copies}}$$

- $\mathbb{Z}^n$ is generated by $e_1 = (1, 0, \ldots, 0)$, ..., $e_n = (0, 0, \ldots, 1)$.

Proposition

Every finitely generated abelian group $\mathbf{G}$*, say* $\mathbf{G} = \langle a_1, \ldots, a_n \rangle$*, is isomorphic to a quotient group of* $\mathbb{Z}^n$*,*

$$\mathbf{G} \simeq \mathbb{Z}^n / \mathbf{H}$$

for some subgroup $\mathbf{H}$*.*

Proof. Class discussion.

- Define a surjective homomorphism $\varphi : \mathbb{Z}^n \to \mathbf{G}$. Use additive notation for the groups here.
- Use the theorem above to give answer.
- What is the kernel like?

Subgroups of Quotient Groups

Proposition

Let $\mathbf{G}$ *be a group and* $\mathbf{N}$ *a normal subgroup. Then*

1. *The subgroups of* $\mathbf{G}/\mathbf{N}$ *are the quotient groups* $\mathbf{K}/\mathbf{N}$*, where* $\mathbf{K}$ *is a subgroup of* $\mathbf{G}$ *containing* $\mathbf{N}$*.*
2. *Moreover,* $\mathbf{K}/\mathbf{N}$ *is a normal subgroup of* $\mathbf{G}/\mathbf{N}$ *if and only if* $\mathbf{K}$ *is a normal subgroup of* $\mathbf{G}$*.*

Third Isomorphism Theorem

Theorem (Third Isomorphism Theorem)

Let $\mathbf{G}$ *be a group and* $\mathbf{N} \subset \mathbf{K}$ *be normal subgroups of* $\mathbf{G}$*. Then*

$$(\mathbf{G}/\mathbf{N})/(\mathbf{K}/\mathbf{N}) \simeq \mathbf{G}/\mathbf{K}.$$

Second Isomorphism Theorem

There is also a so-called second isomorphism theorem:

Theorem

Let $\mathbf{H}$ *and* $\mathbf{K}$ *be subgroups of the group* $\mathbf{G}$*. If* $\mathbf{H}$ *is a normal subgroup, the product* $\mathbf{HK}$ *is a group and there is an isomorphim*

$$\mathbf{K}/\mathbf{K} \cap \mathbf{H} \to \mathbf{HK}/\mathbf{H}.$$

Proof.

- Define a mapping $\pi : \mathbf{K} \to \mathbf{HK}/HH$ by $\pi(k) = k\mathbf{H}$.
- It is easy to verify that π is a group homomorphism: $k\mathbf{H} \cdot k'\mathbf{H} = kk'\mathbf{H}$.
- $\pi(k) = \mathbf{H}$ if and only if $k \in \mathbf{K} \cap \mathbf{H}$.
- Since π is surjective, $\overline{\pi} : \mathbf{K}/\mathbf{K} \cap \mathbf{H}n \to \mathbf{KH}/\mathbf{H}$ is an isomorphism.

Subgroups

$$\begin{tikzcd}
\mathbf{G} \arrow[r, "\phi"] \arrow[d, "\pi"'] & \mathbf{H} \arrow[d, "\pi'"] \\
\mathbf{G}/\mathbf{G}^{(1)} \arrow[r, "\phi'"'] & \mathbf{H}/\mathbf{H}^{(1)}
\end{tikzcd}$$

where π and π' are the natural homomorphisms to the quotient groups.

Discuss: natural, functorial