18.701: Algebra I

Jesse Yang

Fall 2020

§1 September 2, 2020

To begin our discussion, we examine some key examples of groups.

Example (Symmetric Group)
The **symmetric group** $S_n := \{\text{permutations of } \{1, 2, \ldots, n\}\}$

Example (Permutation)
A **permutation** of any set S is a bijection $p : S \to S$

Note: Multiplication of permutations is a composition.

Example (General Linear Group)
The **general linear group** $GL_n(\mathbb{R}) := \{\text{invertible } n \times n \text{ matrices}\}$

Each matrix A defines a function $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^n$, and maps all column vectors $\vec{v} \mapsto A\vec{v}$
Note: Matrix multiplication is a composition.

Example
Let $S = \{1, 2, 3, 4, 5, 6\}$. Let $p : S \to S$ be the permutation:

$$p(1) := 5$$
$$p(2) := 6$$
$$p(3) := 3$$
$$p(4) := 2$$
$$p(5) := 1$$
$$p(6) := 4$$

We can write this permutation in cycle notation: $p = (15)(264)(3)$ which is equivalent to $(15)(264) = (642)(51)$.

Definition (Identity Permutation)
The **identity permutation** maps every element to itself, and can be denoted as 1.

Example
Let $e_1 = \begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix}$, $e_2 = \begin{pmatrix}0\\1\\0\\0\\0\\0\end{pmatrix}$, $\ldots$, $e_6 = \begin{pmatrix}0\\0\\0\\0\\0\\1\end{pmatrix}$. Let P be the associated **permutation matrix**

$$Pe_1 = e_5$$
$$Pe_2 = e_6$$
$$Pe_3 = e_3$$
$$Pe_4 = e_2$$
$$Pe_5 = e_1$$
$$Pe_6 = e_4$$

Then,

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

As an example, we can evalaute in the manner:

$$P\begin{pmatrix}2\\3\\0\\0\\0\\0\end{pmatrix} = P(2e_1 + 3e_2) = 2e_5 + 3e_6 = \begin{pmatrix}0\\0\\0\\0\\2\\3\end{pmatrix}$$

Example

Let q be the **transposition** (25). How can one compute pq? What is $(pq)(2)$?

We can use the composition of permutations to evaluate $(pq)(2) = p(q(2)) = p(5) = 1$.

Definition (sign p)

For $p \in S_n$, sign $p := \det P$. An alternative definition is to write $p = T_1 \cdots T_n$ where each T_i is a transposition. Then sign $p := (-1)^n$. p is **even** if n is even and **odd** if n is odd.

Definition (Group)

A **group** is a set G equipped with a function (equivalently a law of composition or binary operation)

$$G \times G \longrightarrow G$$

$$(a, b) \longmapsto ab$$

satisfying the following axioms:

- **associative**: $(ab)c = a(bc)$ for all $a, b, c \in G$
- **identity**: There exists an element $1 \in G$ such that $1a = a$ and $a1 = a$ for all $a \in G$.
- **inverses**: For every $a \in G$, there exists $b \in G$ (also called a^{-1}) such that $ab = 1$ and $ba = 1$

If only the associative and identity properties are satisfied, the set is a **monoid**.

Given associativity and identity, for any finite sequence $a_1, a_2, \ldots, a_n$ in G, we can define $a_1 a_2 \cdots a_n \in G$ unambiguously. In the case where $n = 0$, the empty product is 1 (the identity of the group).

As a special case for $n \geq 0$, $a^n := a \cdot a \cdots a$ and $a^{-n} := a^{-1} \cdot a^{-1} \cdots a^{-1}$ n times each, so $a^0 = 1$.

Definition (Abelian Group)

If G satisfies the additional axiom

- **commutative**: $ab = ba$ for all $a, b \in G$,

then G is called an **abelian group**.

Some examples of abelian groups include the trivial group, $G = \{1\}$, $(\mathbb{Z}, +)$, S_n for $n \leq 2$, $GL_n(\mathbb{R})$ for $n \leq 1$.

Example

Some examples of abelian groups are:

the trivial group $G = \{1\}$ $(\mathbb{Z}, +)$ S_n for $n \leq 2$ $GL_n(\mathbb{R})$ for $n \leq 1$

Notation: Consider the comparison between multiplicative and additive notation in groups:

multiplicative notation	additive notation
ab	$a + b$
1	0
a^{-1}	$-a$
a^n	na
empty product 1	empty sum 0

Definition (Order)

The **order** of a group is how many elements the group has. For example, $|S_n| = n!$. Groups are referred to as being either **finite** or **infinite**.

Example

Consider the nonabelian group S_3 of order 6.

We let the identity be 1 and let $x := (123)$, meaning $x^2 = (132)$, and thus $x^3 = 1$. Therefore, x generates only these three elements.

Continuing, we can define $y := (12)$, which allows us to obtain $xy = (13)$ and $x^2y = (23)$. Therefore, $1, x, x^2, y, xy, x^2y$ give the whole group S_3.

Proposition 1

S_3 is the largest group generated by two elements x and y satisfying $x^3 = 1$, $y^2 = 1$, $yx = x^2y$. Using notation, we could denote $S_3 \simeq \langle x, y \mid x^3 = 1, y^2 = 1, yx = x^2y \rangle$ where x, y are the **generators**, the equations are **relations**, and the whole expression is a **presentation** of the group.

Proof. Let G be *any* group so generated. A typical element might be generated like

$$x^{-1}x^{-1}yyxy^{-1}$$

and using the relations, $x^{-1} = x^2$ and $y^{-1} = y$ allowing us to rewrite this element as

$$x^2x^2yyxy$$

Continuing, we may use the relation $yx = x^2y$ to obtain

$$x^2x^2x^4yyy = x^8y^3$$

and then we may use relations to reduce the exponents ($x^3 = 1$, $y^2 = 1$):

$$x^2y$$

Thus any element can be reduced in this manner, giving $G = \{x^a y^b \mid 0 \leq a \leq 2, 0 \leq b \leq 1\}$, so $|G| \leq 6$, and we know that such a group of size six exists. We could check the multiplication table of G which would give the same multiplication table as S_3. □

Definition (Subgroup)

Suppose that G is a group and $H \subset G$. Then H is a **subgroup** if it is a group under a binary operation coming from that of G. More explicitly, a subset $H \subset G$ is a subgroup if and only if

- **closure**: if $a, b \in H$, then $ab \in H$.
- **identity**: The 1 of G belongs to H.
- **inverses:** If $a \in H$, then $a^{-1} \in H$.

If the group law of G is written additively, then the subgroup axioms are written equivalently in additive notation.

§2 September 4, 2020

Theorem 2

The subgroups of $\mathbb{Z}$ are the groups $a\mathbb{Z}$ for $a = 0, 1, 2, \ldots$

$$\{0\}, \quad \mathbb{Z}, \quad 2\mathbb{Z}, \quad 3\mathbb{Z}, \quad \ldots$$

Proof. It is easy to check that each of these sets are subgroups.

Consider *any* subgroup $S \subset \mathbb{Z}$. Then we know $0 \in S$.

Case. If $S = \{0\}$, then $S = 0\mathbb{Z}$

Case. If $S \neq \{0\}$, then S contains some nonzero element, say n. Because S is a subgroup, we must also have $-n \in S$, so S must contain a positive integer. Let a be the smallest positive integer in S. We claim that $S = a\mathbb{Z}$.

S contains $\underbrace{a + a + \cdots + a}_{n \text{ times}}$ for any number of terms of n, 0, and $\underbrace{(-a) + (-a) + \cdots + (-a)}_{m \text{ times}}$ for any number of terms m, so $a\mathbb{Z} \subset S$. On the other hand, suppose $n \in S$. Then we can divide to obtain $n = qa + r$, where the remainder r satisfies $0 \leq r < a$. Then $r = n - qa \in S$, but $r < a$, meaning we must have $r = 0$ by the minimality of a.

Therefore $n = qa \in a\mathbb{Z}$. Hence $S \subset a\mathbb{Z}$ too. Since $a\mathbb{Z} \subset S$ and $S \subset a\mathbb{Z}$, we must have $S = a\mathbb{Z}$. □

Remark. $a\mathbb{Z} \subset b\mathbb{Z} \iff a \in b\mathbb{Z} \iff a$ is a multiple of $b \iff b \mid a$.

Fact 3

For $a, b \in \mathbb{Z}$ where a and b are not both 0,

$$\{\text{subgroups of } \mathbb{Z} \text{ containing } a\mathbb{Z}, b\mathbb{Z}\} \longleftrightarrow \{d \geq 1 : d \mid a, d \mid b\}$$

Considering subgroups on the left side, $a\mathbb{Z} + b\mathbb{Z}$ is the smallest such subgroup (it is contained in all others). This corresponds to a common divisor of a, b that is a multiple of all other common divisors, namely $\gcd(a, b)$. In conclusion, we have

$$a\mathbb{Z} + b\mathbb{Z} = \gcd(a, b)\mathbb{Z}$$

Similarly,

$$a\mathbb{Z} \cap b\mathbb{Z} = \operatorname{lcm}(a, b)\mathbb{Z}$$

Corollary (Bézout's Lemma)

Given $a, b \in \mathbb{Z}$, there exists $r, s \in \mathbb{Z}$ such that

$$\gcd(a, b) = ra + sb$$

Definition (Homomorphism)

Consider groups G, H and a function

$$\phi : G \to H$$

Then ϕ is a **homomorphism** if, for all $a, b \in G$,

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

Note: this allows us to multiply in G and then apply the function, or apply the function and then multiply in H. Consider the following examples of homomorphisms:

Example

$$|\,| : \mathbb{C}^\times \to \mathbb{R}^\times \text{ because } |ab| = |a|\,|b| \text{ for all } a, b \in \mathbb{C}^\times$$

$$\det : GL_n(\mathbb{R}) \to \mathbb{R}^\times$$

$$S_n \to GL_n(\mathbb{R}) \text{ or } p \longmapsto \text{ the associated permutation matrix } P$$

$$\text{The composition } S_n \to GL_n(\mathbb{R}) \xrightarrow{\det} \mathbb{R}^\times \text{ where } S_n \to \mathbb{R}^\times \text{ gives the sign}$$

Proposition 4

Given a homomorphism $\phi : G \to H$, then

(a) $\phi(1_G) = 1_H$

(b) $\phi(a^{-1}) = \phi(a)^{-1}$ for all $a \in G$

Proof. $\phi(1 \cdot 1) = \phi(1)\phi(1)$, and $\phi(1 \cdot 1) = \phi(1)$, so $\phi(1) = \phi(1)\phi(1)$. Left multiplying by $\phi(1)^{-1}$, we get

$$1 = \phi(1)$$

Similarly $\phi(a \cdot a^{-1}) = \phi(a)\phi(a^{-1})$ implies

$$\phi(1) = 1 = \phi(a)\phi(a^{-1})$$

so we must have

$$\phi(a^{-1}) = \phi(a)^{-1}$$

□

Definition (Kernel)

Given a homomorphism $\phi : G \to H$, the **kernel** of ϕ is

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

Proposition 5

$\ker \phi$ is a subgroup of G.

Proof. We prove that $\ker \phi$ satisfies the three properties of a subgroup:
Closure: If $a, b \in \ker \phi$, then $\phi(a) = \phi(b) = 1$. Therefore, $\phi(ab) = \phi(a)\phi(b) = 1 \cdot 1 = 1$, so $ab \in \ker \phi$.

Identity: The identity of G, 1_G is in $\ker \phi$ since $\phi(1_G) = 1$.

Inverses: If $a \in \ker \phi$, then $\phi(a) = 1$. Since $1_G \in \ker \phi$, then

$$1 = \phi(1_G) = \phi(a \cdot a^{-1}) = \phi(a) \cdot \phi(a^{-1}) = \phi(a^{-1})$$

Therefore, $a^{-1} \in \ker \phi$ as well. □

Consider the following examples of homomorphisms:

Example

Consider the homomorphism $GL_n(\mathbb{R}) \xrightarrow{\det} \mathbb{R}^\times$. Then,

$$SL_n(\mathbb{R}) := \ker(GL_n(\mathbb{R}) \xrightarrow{\det} \mathbb{R}^\times) = \{A \in M_n(\mathbb{R}) : \det A = 1\}$$

Example

$$A_n := \ker(S_n \xrightarrow{\text{sign}} \{\pm 1\}) \text{ is the } \textbf{alternating group}$$

Example

$$\ker(\mathbb{C}^\times \xrightarrow{||} \mathbb{R}^\times) = \{z \in \mathbb{C}^\times : |z| = 1\} \text{ is known as the } \textbf{circle group}$$

Definition (Image)

Given a homomorphism $\phi : G \to H$, then the **image**

$$\operatorname{im} \phi = \phi(G) = \{\phi(g) : g \in G\}$$

is a subgroup of H.

Example

Given G any group with $x \in G$, then

$$\begin{aligned} \phi : Z &\longrightarrow G \\ k &\longmapsto x^k \end{aligned}$$

is a homomorphism. Its image is

$$\langle x \rangle = \{\dots, x^{-2}, x^{-1}, 1, x, x^2, \dots\} \subset G,$$

the subgroup of G generated by x.

For instance, consider $\langle 2 \rangle$ in $\mathbb{R}^\times$:

$$\{\dots, 2^{-2}, 2^{-1}, 1, 2, 2^2, \dots\}$$

Then consider $\langle i \rangle$ in $\mathbb{C}^\times$:

$$\{\dots, i^{-2}, i^{-1}, 1, i, i^2, i^3, i^4, \dots\} = \{1, i, i^2, i^3\}$$

Definition (Cyclic Group)

A **cyclic group** is a group generated by one element.

The homomorphism $\phi : \mathbb{Z} \longrightarrow G; k \longmapsto x^k$ forms a cyclic subgroup generated by x. Then $\ker(\phi : \mathbb{Z} \longrightarrow G; k \longmapsto x^k)$ is a subgroup of $\mathbb{Z}$. Either $\ker \phi = \{0\}$ or $\ker \phi = n\mathbb{Z}$ for some $n \geq 1$.

Case 1: $\ker \phi = \{0\}$. Then,

- $x^r = x^s \iff x^{r-s} = 1 \iff r - s \in \ker \phi \iff r - s = 0 \iff r = s$
- $\{..., x^{-2}, x^{-1}, 1, x, x^2, \dots\}$ are all distinct.
- $\langle x \rangle$ is an *infinite cyclic group*

Case 2: $\ker \phi = n\mathbb{Z}$ for some $n \geq 1$. Then,

- n is the smallest positive integer such that $x^n = 1$ or "x is of order n"
- $x^r = x^s \iff \cdots \iff r - s \in n\mathbb{Z} \iff r \equiv s \pmod{n}$
- $x^0, x^1, \dots, x^{n-1}$ are distinct and every x^r equals one of these
- $\langle x \rangle$ is a cyclic group of order n

§3 September 9, 2020

Definition (Isomorphism)

An **isomorphism** is a homomorphism $\phi : G \to G'$ that has an inverse homomorphism and has an inverse function ("is bijective"). These definitions are equivalent: if a homomorphism ϕ is bijective, then the inverse function ϕ^{-1} is automatically a homomorphism.

Proof. Given $x, y \in G'$ we must prove $\phi^{-1}(x)\phi^{-1}(y) = \phi^{-1}(xy)$. Let $a = \phi^{-1}(x)$ and $b = \phi^{-1}(y)$. Then

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

meaning $\phi^{-1}(x)\phi^{-1}(y) = ab = \phi^{-1}(xy)$. $G \simeq G'$ means $\exists$ an isomorphism $G \to G'$. $\square$

Definition

We consider the following cases of homomorphisms:

	general case	has inverse
from G to any G'	homomorphism	isomorphism
from G to G	endomorphism	automorphism

Example

$$\begin{aligned} \mathbb{Z}/4\mathbb{Z} &\longrightarrow \langle i \rangle \\ 0 &\longmapsto 1 \\ 1 &\longmapsto i \\ 2 &\longmapsto i^2 \\ 3 &\longmapsto i^3 \end{aligned}$$

is an isomorphism.

In general, each cyclic group is isomorphic to $\mathbb{Z}$ or to $\mathbb{Z}/n\mathbb{Z}$ for some $n \geq 1$.

Example

Let G be any group with $a \in G$. Then,

$$\begin{aligned} \text{inn}_a : G &\longrightarrow G \\ x &\longmapsto axa^{-1} \end{aligned}$$

is an automorphism, called an **inner automorphism** or **conjugation by** a.

Proof. First, $(axa^{-1})(aya^{-1}) = axya^{-1}$, so it is indeed a homomorphism. Furthemore, it has an inverse, namely $\text{inn}_{a^{-1}}$. Because it goes from G to itself and is a homomorphism with an inverse, it must be an automorphism. □

Fact 6

Every automorphism of S_n is inner, except when $n = 6$.

Example

Let $\phi : G \to G'$ be a homomorphism and $K := \ker \phi$. When does b map to the same element as a?

This occurs when

$$\phi(b) = \phi(a) \iff \phi(a)^{-1}\phi(b) = 1 \iff \phi(a^{-1}b) = 1$$
$$\iff a^{-1}b = k \text{ for some } k \in K \iff b \in aK := \{ak : k \in K\}$$

aK is a left coset of K. One consequence of this calculation is that ϕ is injective if and only if $\ker \phi = \{1\}$ (the trivial group).

Definition (Cosets)

Given a group G with $H \subset G$ a subgroup and $a \in G$, then any such

$$aH := \{ah : h \in H\}$$

is called a **left coset** of H. Similarly,

$$Ha := \{ha : h \in H\}$$

is called a **right coset**.

Proposition 7

For subgroup $H \subset G$, the left cosets of H form a **partition** of G.

Proof. Define a relation on G :

$$a \equiv b \text{ means } a = bh \text{ for some } h \in H$$

We claim that

1. $\equiv$ is an equivalence relation
2. The equivalence classes of $\equiv$ are the left cosets of H

Reflexive: Since H, is a subgroup, $1 \in H$, so $a = a \cdot 1$ implies that $a \equiv a$.

Symmetric: Suppose that $a \equiv b$. Then $a = bh$ for some $h \in H$. Right-multiplying by h^{-1} yields $ah^{-1} = b$, and since $h^{-1} \in H$, then $b \equiv a$.

Transitive: Suppose that $a \equiv b$ and $b \equiv c$. Then $a = bh_1$ for some $h_1 \in H$ and $b = ch_2$ for some $h_2 \in H$. Then, substitution yields $a = ch_2h_1$ and $h_2h_1 \in H$ by closure of the subgroup. Therefore, $a \equiv c$.

These three properties indicate that $\equiv$ is an equivalence relation. Then, we consider the equivalence class of $b \in G$:

$$\{a \in G : a \equiv b\} = \{bh : h \in H\} = bH$$

Because equivalence classes form a partition of a set, the left cosets of H form a partition of G. □

Proposition 8

$|aH| = |H|$ for all $a \in G$.

Proof. Consider the bijection

$$\begin{aligned} H &\longrightarrow aH \\ x &\longmapsto ax \end{aligned}$$

the inverse is $y \mapsto a^{-1}y$. □

Definition (Index)
The **index** of H in G is

$$[G : H] = (G : H) := \#\text{ of left cosets of } H \text{ in } G$$

This yields a counting formula:

$$|G| = |H|\,(G : H)$$

The bijection $aH \leftrightarrow Ha^{-1}$ gives that the number of right cosets of H is the same as left.

Theorem 9 (Lagrange's Theorem)
For a finite group G with $H \subset G$ a subgroup, then $|H|$ divides $|G|$.

Corollary
Given G a finite group with $g \in G$, then $\text{ord}(g)$ divides $|G|$.

Proof. Apply Lagrange to $\langle g \rangle$. □

Corollary
If G is a group with $|G| = p$ for prime p then G is cyclic, so $G \simeq \mathbb{Z}/p\mathbb{Z}$ or $G \simeq C_p$ (where C_n is notation for a finite cyclic group of order n).

Proof. Choose $a \in G$ with $a \neq 1$. Then $\text{ord}(a) \mid p$ but $\text{ord}(a) \neq 1$, so $\text{ord}(a) = p$, meaning $\langle a \rangle$ has order p, so $\langle a \rangle = G$. □

§4 September 11, 2020

Corollary (of the Counting Formula)
Given a homomorphism $\phi : G \to G'$, then

$$|G| = |\ker \phi| \, |\operatorname{im} \phi|$$

Proof. Let $K := \ker \phi$. The Counting Formula says

$$|G| = |K| \, (G : K) = |\ker \phi| \, |\operatorname{im} \phi|$$

because each coset of the kernel maps to a single element, i.e., $\{\text{cosets of } K\} \leftrightarrow \operatorname{im} \phi$. □

Definition (Normal Subgroup)
For G a group and $H \subset G$ a subgroup, H is **normal in** G i.e. $H \lhd G$ if

- for all $g \in G$ and $h \in H$, $ghg^{-1} \in H$

Equivalent conditions are

- for all $g \in G$, $gHg^{-1} \subset H$
- for all $g \in G$, $gHg^{-1} = H$
- for all $g \in G$, $\operatorname{inn}_g(H) = H$
- for all $g \in G$, $gH = Hg$

Consider the following examples of normal subgroups:

- the kernel of any homomorphism
- if G is abelian, any subgroup H
- any subgroup of index 2 (The 2 left cosets are H and $G \backslash H$; the 2 right cosets are H and $G \backslash H$)
- The center of G, $Z : \{z \in G : zg = gz \text{ for all } g \in G\}$

Alternative proof ($Z \lhd G$): An isomorphism $\phi : G \to G'$ maps the center of G to the center of G' (by definition of isomorphism), so $\phi(Z) = \phi(Z')$. If we take ϕ to be $\operatorname{inn}_g : G \to G$, so $\operatorname{inn}_g(Z) = Z$ for all $g \in G$ (one of the conditions of being normal).

General principle: Any subgroup that can be described without naming specific elements is normal (it must respect inner automorphisms). For example, the subgroup of G generated by all elements of order 2 is a normal subgroup.

Example (Subgroups of S_3)

$$S_3 = \{1, (12), (13), (23), (123), (132)\}$$

The subgroups of S_3 are $\{1\}$ (order 1), $\langle(12)\rangle$ (order 2), $\langle(13)\rangle$ (order 2), $\langle(23)\rangle$ (order 2), $\langle(123)\rangle$ (order 3), S_3 (order 6). Notice that $\langle(132)\rangle = \langle(123)\rangle$. Any possible subgroup would have to contain one of these subgroups and be contained in S_3, but it would have to have an order that divides 6, and contain other subgroups, so these are the only possibilities.

Proposition 10

Let $\phi : G \to G'$ be a homomorphism.

- If $H \subset G$ is a subgroup, then $\phi(H) \subset G'$ is a subgroup.
- If $H \lhd G$ and ϕ is surjective, then $\phi(H) \lhd G'$

We can also discuss inverse images:

- If $H' \subset G'$ is a subgroup, then $\phi^{-1}(H') \subset G$ is a subgroup ($\phi^{-1}(H') := \{g \in G : \phi(g) \in H'\}$ is not actually a function)
- If $H' \lhd G'$, then $\phi^{-1}(H') \lhd G$

Proof. We only go through the proof of the last statement. Suppose that $g \in G$ and $h \in \phi^{-1}(H')$. Then $\phi(h) \in H'$, so $\phi(ghg^{-1}) = \phi(g)\phi(h)\phi(g)^{-1} \in H'$ since H' is normal in G'. Therefore, $ghg^{-1} \in \phi^{-1}(H')$, meaning $\phi^{-1}(H')$ is normal in G. □

Theorem 11 (Correspondence Theorem)

Given a *surjective* group homomorphism $\phi : G \to G'$, let $K := \ker \phi$. Then

1. There exists a bijective correspondence $\{\text{subgroups of } G \text{ containing } K\} \leftrightarrow \{\text{subgroups of } G'\}$
$$H \longmapsto \phi(H)$$
$$H' \longmapsto \phi^{-1}(H')$$
2. If H corresponds to H', then
   - $H \lhd G \iff H' \lhd G'$
   - $\phi|_H : H \to H'$ is a surjective homomorphism with kernel K.
   - $|H| = |K|\,|H'|$

This proof requires the following facts to be proved:

- $\phi(H)$ is a subgroup of G'
- $\phi^{-1}(H')$ is a subgroup of G containing K
- $\phi^{-1}(\phi(H)) = H$ is an identity
- $\phi(\phi^{-1}(H')) = H'$ is an identity

§5 September 14, 2020

Definition

For any subgroup $H \leq G$ (from here on out we use the $\leq$ notation to denote a subgroup), define

$$G/H := \{\text{left cosets of } H \text{ in } G\}$$

$$H\backslash G := \{\text{right cosets of } H \text{ in } G\}$$

Example (Quotient Group)

Given a normal subgroup $N \lhd G$, how can we make a homomorphism $G \to \cdots$ with kernel N?

Plan: Make a group whose elements *are* the left cosets of N.

- The set is G/N.
- Define the product of aN and bN (elements of G/N) to be

$$aNbN = abNN = (ab)N$$

since N is normal. This shows that the binary operation is defined.

We claim that the set G/N with binary operation $aN, bN \mapsto (ab)N$ is really a group.

Proof. We check

- Associativity:

$$((aN)(bN))\,(cN) = (abN)(cN) = ((ab)c)\,N$$

$$(aN)\,((bN)(cN)) = (aN)(bcN) = (a(bc))\,N$$

which are equal due to the associativity of G.

- Identity:

$$1N = N \in G/N$$

- Inverse of aN is $a^{-1}N$

□

We can also notice that $\pi : G \to G/N$ with $a \mapsto aN$ is a homomorphism (the canonical map from G to G/N), with

$$\ker \pi = \{a \in G : aN = n\} = N$$

As another point of view (for imagining what's going on), the elements of G/N are "the elements of G, except that if $a \equiv b \pmod{N}$, then a and b are considered the same." More precisely, they form "the equivalence classes for $\equiv$" (i.e., the left cosets of N). For instance, we could consider

$$\begin{aligned}\mathbb{Z}/3\mathbb{Z} &= \{3\mathbb{Z}, 1+3\mathbb{Z}, 2+3\mathbb{Z}\} \\ &= \{\overline{0}, \overline{1}, \overline{2}\}\end{aligned}$$

Definition (Coset Representatives)
A set of **coset representatives** for $H \leq G$ is a subset of G consisting of one element in each left coset.

As an example, $\{0, 1, 2\}$ is a set of coset representatives for $3\mathbb{Z} \leq \mathbb{Z}$.

Theorem 12 (First Isomorphism Theorem)
Let $\phi : G \to G'$ be a *surjective* homomorphism, with $N := \ker \phi$. Then,

$$\overline{\phi} : G/N \longrightarrow G'$$

$$aN \longmapsto \phi(a)$$

is an isomorphism.

Proof. We check that $\overline{\phi}$ is defined and injective, i.e., given $a, b \in G$, a, b are in the same coset of N ($aN = bN$) if and only if $\phi(a) = \phi(b)$.

$\overline{\phi}$ is surjective since $G \to G'$ is surjective, i.e. if $\phi(g) = x$ then $\overline{\phi}(gN) = x$.

$\overline{\phi}$ is a homomorphism because

$$\overline{\phi}\left((aN)(bN)\right) = \overline{\phi}(abN) = \phi(ab) = \phi(a)\phi(b) = \overline{\phi}(aN)\overline{\phi}(bN)$$

A bijective homomorphism is an isomorphism. □

Alternative Version of First Isomoprhism Theorem:
For $\phi : G \to G'$ a homomorphism, then $G/\ker\phi \simeq \operatorname{im}\phi$ i.e. it is isomoprhic.

One corollary of this is that if G is finite, $|G| = |\ker\phi|\,|\operatorname{im}\phi|$.

We proceed by beginning our discussion of linear algebra.

Definition (Linear Combination)

Given F a field and V a vector space over F with $S \subset V$, a **linear combination** of elements of S is a sum

$$\sum_{s\in S} a_s s$$

where $a_s \in F$ for each $s \in S$ and $a_s = 0$ for all but finitely many $s \in S$.

Definition (Span, Linearly Independent, Basis)

The **Span** of S is the set of vectors that can be obtained from S by addition and scalar multiplication (finitely many times) i.e. linear combinations of finitely many vectors in S or equivalently, the smallest subspace of V containing S. If S is the empty set, then $\operatorname{Span} S = \{0\}$.

- S **spans** V if $\operatorname{Span} S = V$

  Every vector is a linear combination of vectors in S

- S is **linearly independent** if

$$\sum_{s\in S} a_s s = 0 \implies a_s = 0 \,\forall\, s \in S$$

  Every vector can be expressed as a linear combination in at most one way.

- S is a **basis** of V if S spans V and S is linearly independent

  Each vector can be expressed as a linear combination in *exactly* one way.

Definition (Linear Transformation)

Suppose $T : V \to W$ is a function mapping from one F-vector space to another. Then T is a *homomorphism of vector spaces*, a **linear transformation**, or a *linear map* if

- $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all $v_1, v_2 \in V$
- $T(av) = aT(v)$ for all $a \in F$, $v \in V$

§6 September 16, 2020

As an example of a linear transformation, consider $A \in F^{m \times n}$. Then

$$F^n \xrightarrow{A} F^m$$

$$v \mapsto Av$$

is a linear transformation. Furthermore, every linear transformation $T : F^n \to F^m$ is of this form. Given T, let

$$A = \begin{pmatrix} Te_1 & \cdots & Te_n \end{pmatrix}$$

and then $T(v) = Av$ for all v (by the linear transformation rules).

Definition (Nullspace and Column Space)

We can also consider $\ker A$ (**nullspace**) or $\operatorname{im} A$ (**column space** i.e. span of columns).

As another example, suppose that $S = (v_1, \ldots, v_n)$ where each v_i is a vector in vector space V. Then, we can define a linear transformation: $\phi : F^n \to V$, where

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto \text{“}(v_1 \ldots v_n) \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \text{”} = a_1 v_1 + \cdots + a_n v_n$$

where we may essentially interpret S as a matrix of column vectors. Then,

- S spans $V \iff \phi$ is surjective
- S is linearly independent $\iff \phi$ is injective
- S is a basis $\iff \phi$ is bijective (and then ϕ^{-1} is a linear transformation too, so ϕ is an isomorphism of vector spaces)

Example

Consider $V = \mathbb{R}^2$ and $S = \left(\binom{2}{1}, \binom{1}{3}, \binom{21}{13}\right)$

S spans V, but S is not linearly independent.

Proposition 13

Suppose that S is finite, and S spans V (when such an S exists, V is called **finite-dimensional**).

1. Deleting some vectors in S gives a basis
2. If L is a linearly independent subset of V, inserting some vectors from S gives a basis

Proof of 2. If $\operatorname{Span} L \neq V$, insert some $v \in S$ not in $\operatorname{Span} L$, and repeat; eventually, $\operatorname{Span} L = V$. □

Corollary

Every finite-dimensional vector space has a basis (this is actually true even if V is infinite-dimensional).

Proposition 14

If $F^n \xrightarrow{A} F^m$ is injective, then $n \leq m$.

Proof. Because the linear transformation is injective, $\ker A = \{0\}$, i.e. the system $Av = 0$ has only $v = 0$ as a solution. If $m < n$, then row reduction would show that there exists a nonzero solution, so $m \geq n$. □

Corollary

If $F^n \simeq F^m$, then $n = m$.

Proof. If F^n and F^m are isomorphic, there are bijections (and thus injections) in each direction, so $n \leq m$ and $m \leq n$, implying $n = m$. □

Corollary

Given finite-dimensional V, then every (finite) basis of V has the same number of elements.

Proof. If $F^m \simeq V$ and $F^n \simeq V$ (a basis is isomorphic to the vector space), composing these isomorphisms (one with the inverse of another), $F^m \simeq F^n$, so $m = n$. □

Definition (Dimension)

$$\textbf{dim V} := \#\text{ elements in any basis of } V$$

Definition (Coordinates)

Consider an n-dimensional vector space V with basis $B = (v_1, \ldots, v_n)$ and $w \in V$. Then, there exists a unique

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in F^n$$

such that $x_1 v_1 + \cdots + x_n v_n = w$, where $\vec{x}$ yields the **coordinates** of w with respect to B, i.e. $B\vec{x} = w$. Notice that multiplication by B yields an isomorphism from F^n to V.

We can also discuss *change of bases* (helpful to visualize with a diagram). Suppose $B' = (v_1', \ldots, v_n')$ is another basis. Both B and B' yield isomorphisms from F^n to V, so *defining*

$$P := B^{-1}B'$$

yields a base change matrix from F^n back to F^n. In particular,

$$B' = BP.$$

When taking *coordinates* of w, we can take these coordinates with respect to B ($\vec{x}$) *or* with respect to B' ($\vec{x}'$):

$$P\vec{x}' = \vec{x}.$$

Notice how this is a flipped order from above.

Theorem 15 (Rank-Nullity)

Given a linear transformation $\phi : V \to W$, then

$$\dim(\ker \phi) + \dim(\operatorname{im} \phi) = \dim V$$

where $\dim(\ker \phi)$ is referred to as **nullity** and $\dim(\operatorname{im} \phi)$ is referred to as **rank**.

Notice that this is a vector-space analogue of $|\ker \phi|\,|\operatorname{im} \phi| = |G|$ in a group.

Proof for finite-dimensional V. Choose a basis $(v_1, \ldots, v_k)$ of $\ker \phi$ and extend to a basis

$$(\underbrace{v_1, \ldots, v_k}_{\text{basis of } \ker \phi}, \underbrace{v_{k+1}, \ldots, v_n}_{\text{basis of a space } Z})$$

of V. Then $\ker\phi \oplus Z = V$ (a *direct sum*). We claim that $\phi|_Z : Z \to \operatorname{im}\phi$ is an isomorphism.

We can first check surjectivity (span) by applying ϕ to both sides of the direct sum decomposition, yielding

$$\{0\} \oplus \phi(Z) = \phi(V) \text{ so } \phi(Z) = \operatorname{im}\phi$$

which proves surjectivity.

We can then check injectivity (linearly independent, i.e. $\ker(\phi|_Z) = \{0\}$): If $z \in Z$ and $\phi(z) = 0$, then $z \in \ker\phi$, so $z = 0$ (the only overlap between the kernel and Z, since all vectors in the basis are linearly independent).

Then $\dim(\ker\phi) + \dim Z = \dim V$ (from the direct sum) implies $\dim(\ker\phi) + \dim(\operatorname{im}\phi) = \dim V$. □

§7 September 18, 2020

Example (Matrix of a Linear Transformation)

Given V, W finite-dimensional vector spaces with a linear transformation

$$V \xrightarrow{\phi} W$$

We can choose bases for V (B) and W (C) such that

$$F^n \xrightarrow{\sim} V$$

via B,

$$F^m \xrightarrow{\sim} W$$

via C, and

$$F^n \xrightarrow{A} F^m$$

where $A \in F^{m\times n}$

Theorem 16

Given $\phi : V \to W$, one can choose bases so that

$$A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

where the top left is the $r \times r$ identity matrix and all other elements are 0.

Proof. We let the basis for V be $\underbrace{v_1, \ldots, v_r}_{\text{basis for } Z}, \underbrace{v_{r+1}, \ldots, v_n}_{\text{basis for } \ker \phi}$. Applying ϕ to each of these, we get a basis for W: $\underbrace{w_1, \ldots, w_r}_{\text{basis for } \operatorname{im} \phi}, w_{r+1}, \ldots, w_m$ where for $1 \leq i \leq r$, $w_i = \phi(v_i)$ and we have extended the basis.

We have $\phi(v_1) = w_1$, which means that the first column of A is

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

for m rows. We can continue this structure to establish our desired result for A. □

Corollary

We have that

$$\operatorname{rank}(A) = \operatorname{rank}(A^t)$$

which is equivalent to saying

$$\dim(\text{span of columns}) = \dim(\text{span of rows})$$

Proof. We can reduce to the simple case of A above, in which the row and column space both have dimension r. □

Definition (Linear Operator)

A **linear operator** is a linear transformation from a vector space to itself ($V \to V$), represented by square matrices.

Proposition 17

Suppose that $\dim V < \infty$ (finite-dimensional) and $T : V \to V$ is a linear operator. Then, the following are equivalent:

1. T is injective ($\ker T = \{0\}$)
2. T is surjective ($\operatorname{im} T = V$)
3. T is bijective (T is an isomorphism)
4. $\det T \neq 0$

Proof. Statements 1. and 2. are equivalent from the rank-nullity theorem, since $\dim(\ker T) + \dim(\operatorname{im} T) = \dim V$, so $\dim(\ker T) = 0$ if and only if $\dim(\operatorname{im} T) = \dim V$, implying that $\operatorname{im} T = V$. Thus, the injective and surjective properties are equivalent. Together, these statements are equivalent to bijectivity as well. □

Example (Change of Basis for Operators)

Suppose we have a linear operator $T : V \to V$ with bases B and B' such that

$$F^n \xrightarrow{B} V \quad F^n \xrightarrow{B'} V$$

where basis B yields a matrix A, and basis B' yields a matrix A'.

Letting P be the basechange matrix from B' to B, then

$$A' = P^{-1}AP$$

where it's very helpful to think about this relationship with a diagram.

Definition (Determinant of Linear Operator)

We define the **determinant of a linear operator**

$$\det T := \det A$$

since $\det A$ does not depend on the choice of basis, e.g. another choice leads to

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

Definition (Similarity of Matrices)

Two $n \times n$ matrices A and B are **similar** if

$$B = P^{-1}AP$$

for some invertible $n \times n$ matrix P.

Definition (Eigenvector and Eigenvalue)

Given a finite-dimensional vector space V with an operator $T : V \to V$ and $\lambda \in F$ (possibly 0), an **eigenvector** with **eigenvalue** λ is a (nonzero) vector $v \in V$ such that

$$Tv = \lambda v$$

We call λ an eigenvalue if there exists $v \neq 0$ such that $Tv = \lambda v$. As an example, consider

$$\mathbb{R}^2 \xrightarrow{\begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}} \mathbb{R}^2$$

Then $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector with eigenvalue 5. We could also have $w = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ as an eigenvector with eigenvalue -1. We can write matrix A with respect to the basis (v, w) as a diagonal matrix

$$A = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$$

Proposition 18

For $A \in F^{n \times n}$, A is similar to a diagonal matrix if and only if F^n has a basis of eigenvectors for A.

Explicitly, if $v_1, \ldots, v_n$ is a basis of eigenvectors and $\lambda_1, \ldots, \lambda_n$ are the eigenvalues, letting P be a matrix with $v_1, \ldots, v_n$ as the column vectors and Λ be a diagonal matrix with $\lambda_1, \ldots, \lambda_n$ as entries, the diagram

$$F^n \xrightarrow{P} F^n \xrightarrow{A} F^n = F^n \xrightarrow{\Lambda} F^n \xrightarrow{P} F^n$$

commutes, where

$$A = P\Lambda P^{-1}$$

§8 September 21, 2020

Proposition 19

Given λ, the eigenspace of λ is

$$\{\text{eigenvectors with eigenvalue } \lambda, \text{ including } 0\} = \ker(\lambda I - T) = \ker(T - \lambda I)$$

Proof. To say that v is an eigenvector with eigenvalue λ (or $v = 0$), it is equivalent to say $Tv = \lambda v$, or

$$\lambda v - Tv = 0 \iff (\lambda I - T)v = 0$$

where I is the identity operator. Then, $v \in \ker(\lambda I - T)$, which has the same kernel as its negative operator. □

Proposition 20

The eigenvalues of T are the zeros of $p(t)$ (the characteristic polynomial)

Proof. We know that λ is an eigenvalue if $\ker(\lambda I - T) \neq \{0\}$, which is equivalent to saying (Proposition 41)

$$\det(\lambda I - T) = 0 \iff \det(\lambda I - A)$$

for any matrix A representing T. The determinant yields a polynomial $p(\lambda) = 0$. □

Definition (Characteristic Polynomial)

The **characteristic polynomial** of A (or of T) is

$$p(t) := \det(tI - A)$$

If A happens to be *upper triangular* (below the diagonal is all 0s) then

$$p(t) = (t - a_{11}) \cdots (t - a_{nn})$$

and the eigenvalues of A are the diagonal entries. A numerical example is

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$$

where

$$tI - A = \begin{pmatrix} t-2 & -3 \\ -3 & t-2 \end{pmatrix}$$

and the characteristic polynomial is

$$p(t) = t^2 - 4t - 5 = (t-5)(t+1)$$

so the eigenvalues are 5 and -1. To find the eigenvectors with eigenvalue 5, we can compute

$$\ker(5I - A) = \ker \begin{pmatrix} 3 & -3 \\ -3 & 3 \end{pmatrix}$$

Fact

If $\dim V = n$, then A is $n \times n$, so

$$p(t) = t^n - (\underbrace{a_{11} + \cdots + a_{nn}}_{\text{trace } A})t^{n-1} + \cdots + (-1)^n \det A$$

Notice that $p(0) = \det(-A) = (-1)^n \det A$. If $F = \mathbb{C}$, we can also write

$$p(t) = (t - \lambda_1) \cdots (t - \lambda_n)$$

where λ_i are the eigenvalues (some might be repeated).

Proposition 21

Given nonzero eigenvectors $v_1, \ldots, v_r$ whose eigenvalues $\lambda_1, \ldots, \lambda_r$ are distinct, then $v_1, \ldots, v_r$ are linearly independent.

Proof. We proceed by induction on r. For the base case $r = 0$, the empty set is linearly indepedent.

For $r \geq 1$, suppose $a_1 v_1 + \cdots + a_r v_r = 0$. Applying our linear transformation T to both sides,

$$a_1 \lambda_1 v_1 + \cdots + a_r \lambda_r v_r = 0$$

or we could multiply the whole equation by λ_1:

$$a_1 \lambda_1 v_1 + \cdots + a_r \lambda_1 v_r = 0$$

Subtracting these two equations yields

$$a_2(\lambda_2 - \lambda_1)v_2 + \cdots + a_r(\lambda_r - \lambda_1)v_r = 0$$

By an inductive hypothesis (these $r - 1$ eigenvectors being linearly independent), it must be true that

$$a_i(\lambda_i - \lambda_1) = 0$$

for all $2 \leq i \leq r$. Since the λ are distinct, it must be the case that $a_i = 0$ for all $2 \leq i \leq r$. Plugging these 0 values back into our original hypothesis,

$$a_1 v_1 = 0$$

meaning that $a_1 = 0$. Therefore, the r eigenvectors are linearly independent. □

Corollary

If $p(t)$ has n distinct roots, then

- V has a basis of eigenvectors
- there is a basis with respect to which T is diagonal
- A is **diagonalizable** (similar to a diagonal matrix)

As a *non-example*, consider

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

where

$$p(t) = (t-2)^2$$

with eigenvalues 2,2. The eigenspace of 2 is $\ker(2I - A) = \ker \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$ which is

$$\{\begin{pmatrix} x \\ y \end{pmatrix} : -y = 0\} = \text{Span}\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\}$$

which is not enough to get a basis. Thus, A is *not* diagonizable.

Definition (T-invariant)
Given a linear operator $T : V \to V$ with $W \leq V$ a subspace, then W is **T-invariant** if $TW \subset W$.

In this case, we get an operator $T|_W : W \to W$. If we choose T with respect to a suitable basis (extend a basis of W to basis of V), the matrix of T looks like

$$\begin{pmatrix} T|_W & * \\ 0 & * \end{pmatrix}$$

where $T|_W$ is $r \times r$ if W is r-dimensional.

Definition (Nilpotent Operators)
A linear operator $T : V \to V$ is **nilpotent** if $T^m = 0$ for some m.

For example, if V has basis e_1, e_2, e_3, and

$$e_3 \xrightarrow{T} e_2 \xrightarrow{T} e_1 \xrightarrow{T} 0$$

then T^3 sends every basis to 0, so it sends every linear combination of them to 0, meaning $T^3 = 0$. The matrix of T is

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

As another example, if V has basis $e_1, \ldots, e_5$ and $V = \text{Span}(e_1, e_2, e_3) \oplus \text{Span}(e_4, e_5)$, suppose

$$e_3 \xrightarrow{T} e_2 \xrightarrow{T} e_1 \xrightarrow{T} 0$$

and

$$e_5 \xrightarrow{T} e_4 \xrightarrow{T} 0$$

so T is also a nilpotent operator in this case since $T^3 = 0$. Then, the matrix of T is

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

This matrix has a block structure with two blocks (top 3×3 and bottom 2×2). Then,

$$\ker T = \text{Span}(e_1, e_4)$$

and

$$\dim(\ker T) = \#\text{ chains} = \#\text{ blocks}$$

§9 September 23, 2020

Theorem 22 (Nilpotent Case of Jordan Normal Form)

Let $T : V \to V$ be a linear operator on a finite-dimensional $\mathbb{C}$-vector space (to guarantee that polynomials factor into linear polynomials). Then the following are equivalent

(1) T is nilpotent: $T^m = 0$ for some $m \geq 0$.

(2) All eigenvalues of T are 0.

(3) The characteristic polynomial of T is t^n.

(4) V has a basis consisting of chains of vectors in which T maps each to the next and eventually to $\vec{0}$, where # chains = $\dim(\ker T)$ (each chain has one element which goes directly to 0).

(5) T is represented by a block matrix uniquely determined except for rearranging of blocks, where # blocks = $\dim(\ker T)$

Each similarity class has exactly one nilpotent operator, which corresponds to a unique Jordan normal form (except for permutation of blocks).

Proof. We show that (1) implies (2). Suppose that there is some nonzero eigenvalue, i.e. $Tv = \lambda v$ for some $\lambda \neq 0, v \neq \vec{0}$. Then, $T^2 v = \lambda^2 v \neq 0$, and in general $T^m v \neq 0$ (repeated applying of T, so T^m is never 0, a contradiction of (1).

Next we show that (2) is equivalent to (3). In general,

$$p(t) = (t - \lambda_1) \cdots (t - \lambda_n) = t^n$$

if and only if all eigenvalues are 0.

(4) is equivalent to (5) is relatively easy to show, each set of chains determines block matrices and vice versa.

(2) implies (4).

(1) implies (4). We perform induction on m such that $T^m = 0$. To begin with, consider TV, which is a T-invariant. Additionally,

$$T^{m-1}(TV) = \{0\}$$

By the inductive hypothesis, TV has a basis of chains. Each of the elements in the chain are mapped to by T applied to some v_i. Additionally, some vectors could be mapped directly down to 0 by T. Thus, starting with the basis of TV, insert a preimage v_i of the top vector in each chain for TV. Furthermore, insert vectors k_j in $\ker T$ that together with the bottoms of the chains form a basis for $\ker T$. Letting

$$B := (\text{basis for } TV) \cup \{v_i\} \cup \{k_j\}$$

B is a linearly independent set. Supposing there were a linear dependence, applying T gives a dependence between the basis vectors in TV, so all the non-bottom coefficients are 0. But then there is a dependence between the bottom vectors (the ones that map to 0), but these are a basis of the $\ker T$ and must be independent (leading to a contradiction).

Furthermore, $\#B = \dim(\operatorname{im} T) + \dim(\ker T)$ (each v_i, k_j corresponds with an element of the kernel), so $\#B = \dim(V)$, and since B is linearly independent, it is a basis of V. □

Lemma

Given $T : V \to V$ a finite-dimensional linear operator, then

$$V = V_0 \oplus W$$

where T is nilpotent applied to V_0 and invertible applied to W.

Proof. The dimension can drop, but only finitely many times:

$$V \supset TV \supset T^2V \supset \cdots \supset T^nV$$

so eventually $T^nV = T^{n+1}V = \cdots$. Defining $V_0 := \ker(T^n)$ and $W := T^nV$, then $T|_{V_0}$ is nilpotent and $W = TW$ (bijective), so $T|_W$ is invertible. We check

$$V_0 \cap W = \ker(T^n) \cap W = \{w \in W : T^n w = 0\} = \{\vec{0}\}$$

since $T|_W$ is injective. Then, we can check

$$\dim(\ker T^n) + \dim(\operatorname{im} T^n) = \dim V$$

by the rank-nullity theorem, and $\dim(\ker T^n) = \dim V_0$ and $\dim(\operatorname{im} T^n) = \dim W$, so $V_0 \oplus W$ must be all of V. □

Definition (Jordan Block)

A **Jordan block** is a block of equal eigenvalues along the diagonal, with 1 above the upper diagonal.

A matrix is diagonalizable if all Jordan blocks are 1×1.

Theorem 23 (Jordan Normal Form)

The linear operator $T : V \to V$ in a finite-dimensional $\mathbb{C}$-vector space is represented by a matrix with Jordan blocks along the diagonal, unique except for rearrangment of the blocks.

The diagonal entries are the eigenvalues. An equivalent form of this theorem is that any $A \in \mathbb{C}^{n \times n}$ is similar to a unique matrix in Jordan normal form, except with respect to rearrangement of blocks.

Proof. We proceed by induction on $\dim V$. Suppose $\dim V \geq 1$. Let $\lambda \in C$ be a zero of the characteristic polynomial $p(t)$ (this works because we are working over $\mathbb{C}$), so λ is an eigenvalue. Replacing T by $T - \lambda I$ to assume 0 is an eigenvalue (then if we get a Jordan normal form with eigenvalue 0, we just add back the λ afterwards). Then, consider

$$V = V_0 \oplus W$$

and $V_0 \neq 0$ (there is some eigenvector with eigenvalue 0) so $\dim W < \dim V$, meaning $T|_W$ has a Jordan normal form by the inductive hypothesis. Furthermore $T|_{V_0}$ has a Jordan normal form by the nilpotent version of the theorem, and then you can stick these Jordan normal forms together in a matrix since $V = V_0 \oplus W$. □

The number of blocks is determined by $\dim \ker(T - \lambda I)$ (total blocks), $\dim \ker((T - \lambda I)^2)$ (total blocks with size ≥ 2), and so on...

Definition

$v \in V$ is a **generalized eigenvector** with eigenvalue λ if

$$(T - \lambda I)^m v = 0$$

for some m. We can also let

$$V_\lambda := \{\text{generalized eigenvectors with eigenvalue } \lambda\}$$

Corollary

If

$$p(t) = (t - \lambda_1)^{e_1} \cdots (t - \lambda_r)^{e_r}$$

with λ_i distinct, then

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$$

with dimension

$$n = e_1 + \cdots + e_r$$

where

$$e_i = \#\text{ diagonal entries equal to } \lambda_i = \sum \text{sizes of blocks with diagonal } \lambda_i$$

Also,

$$\dim(\underbrace{\ker(T - \lambda I)}_{\text{eigenspace of } \lambda}) = \#\text{ blocks with diagonal } \lambda$$

§10 September 25, 2020

Definition (Dot Product)

Given column vectors, $x, y \in \mathbb{R}^n$

$$x \cdot y := x^t y = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \cdots + x_n y_n \in \mathbb{R}$$

Definition (Length)

$|x| := \sqrt{x \cdot x} \in \mathbb{R}_{\geq 0}$. For nonzero vectors x and y forming an angle θ,

$$x \cdot y = |x|\,|y| \cos\theta$$

Definition (Orthogonality)

$x \perp y$ means $x \cdot y = 0$ (either x or y is $\vec{0}$, or else $\theta = \pi/2$.

Definition (Orthonormal Basis)

A sequence of vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$ is an **orthonormal basis** if $|v_i| = 1$ for all i and $v_i \perp v_j$ for all pairs $i \neq j$. We have

$$v_i \cdot v_j = \delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Theorem 24 (Orthogonal Matrices)

The following are equivalent for $A \in \mathbb{R}^{n \times n}$:

(1) $|Ax| = |x|$ for all $x \in \mathbb{R}^n$ (A preserves length)

(2) $(Ax) \cdot (Ay) = x \cdot y$ for all $x, y \in \mathbb{R}^n$

(3) $A^t A = I$

(4) The columns of A form an orthnormal basis

A is called an **orthogonal matrix**.

Proof. Given (1), we prove (2) by expressing $x \cdot y = \frac{|x+y|^2 - |x|^2 - |y|^2}{2}$. (2) implies (1) because $|x| = \sqrt{x \cdot x}$.

(2) yields $x^t A^t A y = x^t y$ for all x, y (in particular, we can consider $x = e_i, y = e_j$ which yields $(A^t A)_{ij} = \delta_{ij}$ for all i, j, meaning $A^t A = I$, which is equivalent to saying the ith row of A^t (ith column of A) dotted with the jth column of A is δ_{ij} for all i, j (meaning the columns form an orthonormal basis. □

Definition (Orthogonal Group)

The **orthogonal group** is

$$O_n := \{\text{orthogonal } n \times n \text{ matrices}\} \subset GL_n(\mathbb{R})$$

O_n is a subgroup of $GL_n(\mathbb{R})$, which can be shown primarily using property (3) above of orthogonal matrices.

Proposition 25 (Determinant of Orthogonal Matrices)

If $A \in O_n$, then $\det A \in \{\pm 1\}$. For each $n \geq 1$, both values are possible.

Proof. Given $A \in O_n$, then $A^t A = I$, so

$$(\det A)(\det A) = 1$$

meaning $\det A = \pm 1$. □

Definition (Special Orthogonal Group)
The **special orthogonal group** is

$$SO_n := \ker(O_n \xrightarrow{\det} \{\pm 1\}) = \{\text{orthogonal matrices of } \det 1\}$$

and the index

$$(O_n : SO_n) = 2$$

Considering $n = 2$, we look at the possible orthonormal bases of $\mathbb{R}^2$. The first vector can be expressed as

$$\begin{pmatrix} c \\ s \end{pmatrix}$$

where $c = \cos\theta$ and $s = \sin\theta$ for some angle $\theta \in \mathbb{R}$. Then, our second vector can be either 90 degrees clockwise or counterclockwise, i.e.

$$\begin{pmatrix} -s \\ c \end{pmatrix} \text{ or } \begin{pmatrix} s \\ -c \end{pmatrix}$$

Corollary

$$O_2 = \{\underbrace{\begin{pmatrix} c & -s \\ s & c \end{pmatrix}}_{\det 1}, \underbrace{\begin{pmatrix} c & s \\ s & -c \end{pmatrix}}_{\det -1} : c = \cos\theta, s = \sin\theta\}$$

In particular,

$$SO_2 = \{\begin{pmatrix} c & -s \\ s & c \end{pmatrix}\} = \{\text{rotations around the origin}\} \simeq \{z \in \mathbb{C} : |z| = 1\}$$

while (call the other set S for now)

$$S := \text{the nontrivial coset of } SO_2 \text{ in } O_2 = \{\begin{pmatrix} c & s \\ s & -c \end{pmatrix}\}$$

Proposition 26

If $A \in S$, then A is a reflection in some line L through $\vec{0}$ (at angle $\theta/2$).

Proof. The eigenvalues are 1 and -1, so we can consider unit eigenvectors v, w. We claim that $v \perp w$. In particular, consider

$$v \cdot (-w) = (Av) \cdot (Aw) = v \cdot w$$

where the first equation stems from the eigenvalues and the second stems from the orthogonal matrix A, meaning

$$-(v \cdot w) = v \cdot w$$

so $v \cdot w$, meaning $v \perp w$. □

Consider now the case where $n = 3$. Start with a unit vector $u \in \mathbb{R}^3$ and an angle $\theta \in \mathbb{R}$. Let

$$u^\perp := \{v \in \mathbb{R}^3 : v \perp u\}$$

be a plane. When your right thumb is in the direction of u, your fingers point in the direction of θ (the rotation) in $u^\perp$.

Definition (Rotation in 3D Space)

$$\underbrace{\rho_{(u,\theta)}}_{\text{spin}} := \text{the linear operator } \rho : \mathbb{R}^3 \to \mathbb{R}^3$$

such that

$$\rho(u) = u$$

- $\rho|_{u^\perp}$ is a counterclockwise rotation by θ (as viewed from u)

This is a rotation about $\vec{0}$ in $\mathbb{R}^3$ and corresponds to a 3×3 matrix.

Theorem 27

$$\{3 \times 3 \text{ rotation matrices}\} = SO_3$$

This says that the composite of a rotation about one axis and a rotation about another is equivalent to one single rotation about an axis.

Proof. Let $\rho_{(u,\theta)}$ be a rotation. Choose an orthonormal basis $(u,v,w) = P \in O_3$. We can perform a change of basis by P, which yields

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{pmatrix}$$

since u is held constant and the rest is rotated, meaning

$$\rho = P \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{pmatrix}}_{\text{in } SO_3} P^{-1}$$

Since $SO_3 \lhd O_3$, $\rho \in SO_3$ too. Now suppose $A \in SO_3$, Then $A^tA = I$ and $\det A = 1$. We claim that 1 is an eigenvalue of A. Consider

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

$$= \det(I-A) = (-1)^3 \det(A-I) = -\det(A-I)$$

so $\det(A-I) = 0$, so 1 is an eigenvalue of A. Then, let u be an eigenvector with eigenvalue 1, i.e. u is a fixed vector. We can also scale (multiply by a scalar) to assume $|u| = 1$. By an orthogonal change of basis, without loss of generality,

$$u = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

meaning

$$u^\perp = \{ \begin{pmatrix} 0 \\ * \\ * \end{pmatrix} \} \text{ is } A - \text{invariant},$$

so

$$A = \begin{pmatrix} 1 & & \\ & * & * \\ & * & * \end{pmatrix}$$

where the bottom 2×2 matrix must also be orthogonal and have determinant 1. This matrix must be in SO_2 and therefore ought to be a rotation matrix, so A is a rotation. $\square$

§11 September 28, 2020

Definition (Isometry)

$f : \mathbb{R}^n \to \mathbb{R}^n$ is called an **isometry** if it preserves distances:

$$|f(u) - f(v)| = |u - v| \text{ for all } u, v \in \mathbb{R}^n$$

As examples, consider

- "orthogonal linear operators" preserve dot products and thus distances

$$\mathbb{R}^n \xrightarrow{A} \mathbb{R}^n, x \mapsto Ax \text{ for each } A \in O_n$$

- "translation by b"

$$\mathbb{R}^n \xrightarrow{t_b} \mathbb{R}^n, x \mapsto x + b \text{ for } b \in \mathbb{R}^n$$

Theorem 28

Every isometry $f : \mathbb{R}^n \to \mathbb{R}^n$ is $t_b \circ A$ for some unique $b \in \mathbb{R}^n$ and $A \in O_n$.

Lemma

An isometry $f : \mathbb{R}^n \to \mathbb{R}^n$ fixing $\vec{0}$ is a linear operator

Proof. We can express dot products in terms of $\vec{0}$ and distances:

$$u \cdot v = \frac{|u-0|^2 + |v-0|^2 - |u-v|^2}{2}$$

so f preserves dot products because it preserves distances.

We can also express a sum in terms of dot products:

$$z = x + y \iff (z-x-y)\cdot(z-x-y) = 0 \iff z\cdot z - 2x \cdot z - \ldots = 0$$

so f preserves sums because it preserves all the dot products.

Similarly, f preserves scalar multiplication by each $c \in \mathbb{R}$. □

Proof of Theorem. Let $b = f(0)$. Then $t_b^{-1} f$ is an isometry mapping 0 to 0, so

$$t_b^{-1} f = A \text{ for some orthogonal matrix } A$$

meaning

$$f = t_b \circ A$$

If $f = t_b A$, then

$$f(0) = t_b(A(0)) = t_b(0) = b$$

so b is determined by f, and then

$$A = t_b^{-1} f$$

is determined too. □

Definition

Though not standard notation, we refer to

$$M_n := \{\text{isometries } \mathbb{R}^n \to \mathbb{R}^n\} = \{t_b A\}$$

These refer to the maps

$$\{x \mapsto Ax + b \mid A \in O_n, b \in \mathbb{R}^n\}$$

This is a subgroup of the group of permutations $\mathbb{R}^n$.

$$\mathbb{R}^n \simeq \{\text{translations}\} \text{ and } \{\text{orthogonal operators}\} = O_n$$

where translations are a normal subgroup and orthogonal operators are a subgroup. M_n is generated by these two subgroups, i.e.,

$$\begin{aligned} \mathbb{R}^n \times O_n &\to M_n \\ t_b, A &\mapsto t_b A \end{aligned}$$

is a bijection (of sets). We consider whether this is an isomorphism of groups, i.e., whether or not

$$(t_b t_{b'})(AA)' = (t_b A)(t_{b'} A')$$

which is true only if

$$t_{b'} A = A t_{b'}$$

which is not the case for all A and b'. Instead,

$$A t_b = t_{A(b)} A$$

Example

There exists a homomorphism

$$\begin{aligned} M_n &\xrightarrow{\pi} O_n \\ (x \mapsto Ax + b) &\mapsto A \end{aligned}$$

In this case,

$$\ker \pi = \{\text{translations}\} = \mathbb{R}^n$$

Example (Change of Origin)

Given an isometry

$$\mathbb{R}^n \xrightarrow{f} \mathbb{R}^n$$

we can use a translation map $\mathbb{R}^n \xrightarrow{t_a} \mathbb{R}^n$ (a is the new origin), so our new isometry can be expressed as

$$f' = t_a^{-1} f t_a$$

Definition (Orientation-Preserving)

An isometry $x \mapsto Ax + b$ is **orientation-preserving** if $\det A = 1$ (otherwise it is known as **orientation-reversing** if $\det A = -1$).

Theorem 29 (Isometries in $\mathbb{R}^2$)

Every isometry of $\mathbb{R}^2$ is one of the following:

- translation ($x \mapsto x + b$)
- rotation around some point
- reflection r_l across any line l
- glide reflection ($t_b r_l$) for some l and some nonzero b parallel to l

The first two are orientation-preserving while the latter two or orientation-reversing.

Proof. Let f be an isometry mapping $x \mapsto Ax + b$.

Orientation-preserving Case

We must have $A \in SO_2$. If $A = I$, we have a translation. We claim that if $A \neq I$, then f has a fixed point. This means we need to solve $Ax + b = x$ or $(A - I)x = -b$. A is a non-trivial rotation and thus does not fix any nonzero vector, so $\ker(A - I) = \{0\}$. Thus, $A - I$ is invertible, so $(A - I)x = -b$ has a solution.

We can change the origin to assume that the fixed point is 0. Then, f is an orthogonal matrix and must be a rotation.

Orientation-reversing Case

$f = t_b A$ where A is an element of O_2 not in SO_2, i.e., a reflection in a line L through the origin. We change the origin to $\frac{b}{2}$:

$$t_{b/2}^{-1} f t_{b/2} = t_{-b/2} t_b A t_{b/2} = t_{b/2} t_{A(b/2)} A = t_m A$$

where

$$m = \frac{b + Ab}{2}$$

Since Ab is the reflection of b over L, m is on line L. If $m = 0$, we have a reflection and if $m \neq 0$, we have a glide reflection. □

§12 September 30, 2020

Definition (Discrete Subgroups of $\mathbb{R}$)

A subgroup $G \leq \mathbb{R}$ is **discrete** if there exists $\varepsilon > 0$ such that every nonzero $g \in G$ satisfies $|g| \geq \varepsilon$.

Then, if $g, h \in G$ are distinct, $|g - h| > \varepsilon$.

Theorem 30

Every discrete subgroup $G \leq \mathbb{R}$ is

$$\{0\} \text{ or } \mathbb{Z}a \text{ for some } a \in \mathbb{R}_{>0}$$

Proof Idea. Assuming $G \neq \{0\}$. There is some smallest positive element a of G (there is a nonzero element and thus a positive element $\geq \varepsilon$). If there were any other element, we could construct an element smaller than a. □

Example (Finite Subgroups of O_2)

Given $x :=$ rotation by $\frac{2\pi}{n}$ about O and $y :=$ reflection in a line l through O,

$$C_n := \langle x \rangle \text{ is cyclic of order } n$$

$$D_n := \text{ subgroup of } O_2 \text{ generated by } x, y \text{ is the dihedral group of order } 2n$$

$$= \langle x, y \mid x^n = 1, y^2 = 1, yx = x^{-1}y \rangle = \{\underbrace{1, x, x^2, \ldots, x^{n-1}}_{\text{rotations}}, \underbrace{y, xy, \ldots, x^{n-1}y}_{\text{reflections of order 2}}\}$$

$C_n \triangleleft D_n$ with index 2.

For small cases, we have

$$D_1 \simeq C_2$$
$$D_2 \simeq C_2 \times C_2$$
$$D_3 \simeq S_3$$

For $n \geq 3$,

$$D_n = \{\text{symmetries of } P\} := \{\text{isometries } f : \mathbb{R}^2 \to \mathbb{R}^2 \text{ mapping } P \to P\}$$

where P is a regular n-gon.

Theorem 31

Every finite subgroup $H \leq SO_2$ is C_n for some $n \geq 1$.

Proof. Let $S = \{\theta \in \mathbb{R} : \rho_\theta \in H\} \leq \mathbb{R}$, where ρ_θ is the rotation by θ around $\vec{0}$. Since H is a finite subgroup, S is *discrete* (finitely many elements in any bounded interval). Then,

$$S = \mathbb{Z}a \text{ for some } a \in \mathbb{R}_{>0}$$

Also $2\pi \in S$, so $2\pi = na$ for some $n \in \mathbb{Z}_{>0}$. Then,

$$a = \frac{2\pi}{n} \text{ and } S = \mathbb{Z} \cdot (\frac{2\pi}{n})$$

meaning

$$H = \{\text{rotations by multiples of } \frac{2\pi}{n}\} = C_n.$$

□

Theorem 32

Every finite subgroup $G \leq O_2$ is C_n or D_n for some $n \geq 1$.

Case 1: $G \subset SO_2$, meaning $G = C_n$ for some n, by the previous theorem.

Case 2: $G \not\subset SO_2$. Considering $O_2 \xrightarrow{\det} \{\pm 1\}$ with kernel SO_2, we instead consider

$$H = \ker(G \to \{\pm 1\})$$

Then $G \to \{\pm 1\}$ has two nonempty fibers, namely $H \subset SO_2$ and $Hr \subset O_2 \backslash SO_2$ for some reflection r. Since $H \subset SO_2$, $H = C_n$ for some n. Then G is generated by C_n and a reflection r, so $G \simeq D_n$.

It turns out that discrete subgroups of O_2 are the same as the finite subgroups of O_2.

§13 October 2, 2020

Definition (Discrete subsets of $\mathbb{R}^n$)

Let S be a subset of $\mathbb{R}^n$. Call $s \in S$ **isolated** if some open ball around s contains no other points of S. Call S **discrete** if every point of S is isolated.

Fact

A subgroup $G \leq \mathbb{R}^n$ is discrete if and only if 0 is isolated in G.

Theorem 33 (Discrete subgroups of $\mathbb{R}^2$)

The discrete subgroups of $\mathbb{R}^2$ are

- $\{0\}$
- $\mathbb{Z}a$, for any nonzero $a \in \mathbb{R}^2$
- $\mathbb{Z}a + \mathbb{Z}b$, for any linearly independent $a, b \in \mathbb{R}^2$

Proof. Let $G \leq \mathbb{R}^2$ be a discrete subgroup. If $G = \{0\}$, we are done. Otherwise, choose a one-dimensional subspace (a line) L such that $G \cap L \neq \{0\}$. Then, $G \cap L = \mathbb{Z}a$ for some $a \in L$ (which follows from the one-dimensional case). If $G = \mathbb{Z}a$, we are also done.

Otherwise, there is some other element $b \in G \backslash \mathbb{Z}a$ that is closest to L.

Lemma

Any bounded subset B of $\mathbb{R}^2$ contains only finitely many elements of G.

Proof. We can split our bounded subset B up with a grid with squares that are bounded by ε. Then, there can be no more than two points in the same square, meaning we cannot possibly have infinitely many elements. □

We are guaranteed to be able to chosen such a b since G is discrete, and we can shift elements by multiples of a into a bounded region (parallelogram formed by a and b). Then

$$G = \mathbb{Z}a + \mathbb{Z}b$$

since any other element would be able to be shifted to be closer to L than b. □

Theorem 34

Every finite subgroup $G \leq M_2$ is isomorphic to C_n or D_n for some $n \geq 1$.

Proof. First, we show that there is a finite set S such that $gS = S$ for every $g \in G$. Choose $s \in \mathbb{R}^2$. We form the *orbit*

$$Gs = \{g's : g' \in G\}$$

also known as the G-orbit of s. Then,

$$g(Gs) = (gG)s = Gs$$

so we can take Gs to be S. Now we show that there is a unique point x minimizing

$$|x - s_1|^2 + \cdots + |x - s_n|^2$$

where s_i are the elements of the orbit S. Then, our sum is

$$(x - s_1) \cdot (x - s_1) + \cdots + (x - s_n)(x - s_n) = n\left(x - \frac{s_1 + \cdots + s_n}{n}\right)^2 + \text{ constant}$$

where the constant depends on the s_i but not on x. Thus, our expression is minimized when

$$x = \frac{s_1 + \cdots + s_n}{n}$$

which is also the centroid of S. Then, x is a fixed point, i.e., $gx = x$ for all $g \in G$. This is because g preserves S and preserves distances, so g must preserve x. Then, we let x be the new origin, so G is a group of isometries fixing the origin, so $G \leq O_2$. The finite subgroups of O_2 are C_n or D_n so $G \simeq C_n$ or D_n. □

Definition (Discrete Subgroup in $\mathbb{R}^n$)

Considering a subgroup $G \leq M_n$, view

$$G \leq M_n \longleftrightarrow \{(A, b) : A \in O_n, b \in \mathbb{R}^n\} \subset \mathbb{R}^{n^2+n}$$

We call G **discrete** if and only if G is discrete as a subset of $\mathbb{R}^{n^2+n}$, which is equivalent to saying there does not exist a sequence $g_1, g_2, \ldots$ of distinct elements of G converging to 1.

We know there is a homomorphism $\pi : M_2 \to O_2$ with $\ker \pi = \mathbb{R}^2$ being the translations. For a subgroup $G \leq M_2$,

$$\pi|_G : G \to \overline{G} := \pi(G)$$

Then we can let L be the set of {translations in G} which is $\ker(\pi|_G)$ and a subgroup of $\mathbb{R}^2$. $\overline{G}$ is the *point group* of G while L is the *translation group* of G, so

$$L \to G \to \overline{G}$$

Proposition 35

Each $A \in \overline{G}$ maps L to L.

Proof. Suppose $A \in \overline{G}$ and $t_a \in L$. We need to show $t_{Aa} \in L$. By definition, A is the image of some $t_b A \in G$. Then

$$(t_b A) t_a (t_b A)^{-1} \in G$$

so

$$t_b A t_a A^{-1} t_b^{-1} = t_b t_{Aa} A A^{-1} t_{-b} = t_{b+Aa-b} = t_{Aa} \in G$$

Thus, $t_{Aa} \in L$. □

§14 October 5, 2020

Theorem 36

Given a discrete subgroup $G \leq M_2$,

(1) L is a discrete subgroup of $\mathbb{R}^2$

(2) If $L \neq \{0\}$, then $\overline{G} = C_n$ or D_n, where n is $1, 2, 3, 4$, or 6

(3) If $L = \{0\}$, then G is finite

Proof. (1) $L \subset G$ and a subset of a discrete set is also discrete.

(2) Let a be the shortest nonzero vector in L. If ρ is a rotation by θ in $\overline{G}$, then $\rho a \in L$. Then $\theta \geq \frac{2\pi}{6}$ since otherwise, $\rho a - a \in L$ is shorter than a. Thus, $\overline{G}$ is a discrete subgroup of O_2, so $\overline{G} = C_n$ or D_n and $n \leq 6$. Also, if $n = 5$, then

$$a + \rho_{2(\frac{2\pi}{5})} a$$

is shorter than a, which is a contradiction.

(3) Suppose $L = \{0\}$, meaning there are no nonzero translations in G and G maps isomorphically to $\overline{G}$. Let H be the subgroup of G consisting of the orientation-preserving isometries in G. If H has nontrivial rotations around two different points, then H contains a nontrivial translation, which is a contradiction. Thus, after changing the origin, H is a discrete subgroup of SO_2, so $H \simeq C_n$ for some n. Then G is finite: $|G| \leq 2|H|$ (orientation-preserving or orientation-reversing subgroups), so $G \simeq C_n$ or D_n. □

Example

Each $A \in GL_n$ gives a bijection $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^n$. We can package them all into one function

$$\begin{aligned} GL_n \times \mathbb{R}^n &\to \mathbb{R}^n \\ (A, \vec{x}) &\mapsto A\vec{x} \end{aligned}$$

GL_n is said to act on $\mathbb{R}^n$, and similarly, S_n acts on $\{1, 2, \ldots, n\}$, M_2 acts on $\mathbb{R}^2$.

Definition (Operations/Actions)

Given a group G and set S, an **operation** (**action**) of G on S is a function

$$\begin{aligned} G \times S &\to S \\ (g, s) &\mapsto gs \end{aligned}$$

such that

(1) $1s = s$ for all $s \in S$

(2) $(gh)s = g(hs)$ for all $g, h \in G$ and $s \in S$

If we fix $g \in G$, we get a permutation of S:

$$\begin{aligned} m_g : S &\to S \\ s &\mapsto gs \end{aligned}$$

with inverse $m_{g^{-1}}$. The map

$$\begin{aligned} G &\to \mathrm{Perm}(S) \\ g &\mapsto m_g \end{aligned}$$

is a group homomorphism:

$$m_g m_h = m_{gh}$$

because

$$m_g(m_h(s)) = g(hs) = (gh)s = m_{gh}(s)$$

for all $s \in S$.

Giving an operation of G on S is the same as giving a homomorphism $G \to \mathrm{Perm}(S)$.

Definition (Faithful Operation)

An operation is **faithful** if the only $g \in G$ such that $m_g = 1$ is $g = 1$, i.e. $G \to \mathrm{Perm}(S)$ is injective.

Definition (Orbit and Stabilizer)

Suppose G acts on S and $s \in S$. Then, the **orbit** of s is

$$O_s = Gs := \{gs : g \in G\}$$

which is a subset of S. The **stabilizer** of s is

$$G_s = \text{Stab}_G(s) := \{g \in G : gs = s\}$$

which is a subgroup of G.

Example

Consider the group S_n which acts on $\{1, 2, \ldots, n\}$.

- The orbit of n is the whole set $\{1, 2, \ldots, n\}$ (transitive action: one orbit)
- The stabilizer of n is S_{n-1}
- $\ker(S_n \to \text{Perm}(\{1, 2, \ldots, n\})) = \{1\}$

Proposition 37

The orbits form a partition of S.

Proof. Define $s' \sim s$ if $s' = gs$ for some $g \in G$. Then, each orbit is an equivalence class, meaning they must partition S. □

Proposition 38

Let G act on S with $s \in S$ and let $H := \text{Stab}(S)$ and O_s be the orbit of s. Then, there exists a bijection

$$\begin{aligned} G/H &\xrightarrow{\epsilon} O_s \\ gH &\mapsto gs \end{aligned}$$

Proof. For $g, \gamma \in G$,

$$\gamma s = gs \iff g^{-1}\gamma s = s \iff g^{-1}\gamma \in H \iff \gamma \in gH$$

□

Corollary

$$|O_s| = (G : \mathrm{Stab}(s))$$
$$|O_s|\,|\mathrm{Stab}(s)| = |G|$$

For any $H \leq G$, normal or not, G acts on G/H: $g \in G$ maps a left coset $C \in G/H$ to gC (which is transitive). Furthermore, the stabilizer of $H \in G/H$ is H.

§15 October 7, 2020

Example

Suppose that $s' = as$, where $s \in S$ (set) and $a \in G$ (group). We consider how the stabilizers are related.

If $g \in \mathrm{Stab}(s)$, then $aga^{-1} \in \mathrm{Stab}(s')$, because it will map s' back to itself. The conclusion then is that $\mathrm{Stab}(s') = a\,\mathrm{Stab}(s)a^{-1}$, a conjugate group.

Theorem 39 (Finite Subgroups of SO_3)

The finite subgroups of SO_3 ($\{$rotations in $\mathbb{R}^3$ fixing $\vec{0}\}$) are

Grp.	Name	Order	Description	Stabs.	Orbits
C_n	cyclic	n	generated by $\rho_{(u,2\pi/n)}$	n, n	$1, 1$
D_n	dihedral	$2n$	gen. by $\rho_{(u,2\pi/n)}, \rho_{(v,\pi)}$, $u \perp v$	$2, 2, n$	$n, n, 2$
T	tetrahedral	12	rot. symms. of a tetrahedron	$2, 3, 3$	$6, 4, 4$
O	octahedral	24	rot. symms. of an octahedron	$2, 3, 4$	$12, 8, 6$
I	icosahedral	60	rot. symms. of an icosahedron	$2, 3, 5$	$30, 20, 12$

Let G be a finite subgroup of SO_3. For $g \in G$ with $g \neq 1$, g will be a rotation and will fix two unit vectors (along the axis of rotation), i.e., the *poles* of g. Let

$$P := \cup_{g \neq 1}\{\text{poles of } g\}$$

For example, if $G = C_n$, $|P| = 2$ (rotational axis contributes two poles), and if $G = D_n$, $|P| = 2 + 2n$ (rotational axis and each line of reflection contributes two poles).

Lemma

If $p \in P$ and $g \in G$, then $gp \in P$.

Proof. Since p is a pole, $\text{Stab}(p) \neq \{1\}$ (bigger than just identity). Then,

$$\text{Stab}(gp) = g\,\text{Stab}(p)g^{-1} \neq \{1\}$$

since the stabilizers have the same size, so gp is a pole. This indicates that G acts on P. □

Let $N = |G|$. Suppose that P is composed of k orbits, where orbit i has n_i elements. Consider p_1 in orbit 1. Then, let $r_1 := |\text{Stab}(p_1)|$, and define r_i similarly. Then, by the orbit-stabilizer theorem,

$$n_i r_i = N \text{ and } 1 < r_i \leq N \text{ (since it is at least a pole)}$$

We consider pairs $(g, p) : g \neq 1$ is a rotation and p is a pole of g. We count the number of elements of $\{(g, p)\}$. We can count this in two ways, first the number of poles per g:

$$\sum_{g \in G, g \neq 1} 2 = 2(N-1) = 2N - 2$$

We could also count the number of rotations per pole:

$$\sum_{p \in P} (|\text{Stab}(p)| - 1) = n_1(r_1 - 1) + \cdots + n_k(r_k - 1) = \frac{N}{r_1}(r_1 - 1) + \cdots + \frac{N}{r_k}(r_k - 1)$$

where we have grouped the terms based on the orbit they are in. Then

$$\frac{N}{r_1}(r_1 - 1) + \cdots + \frac{N}{r_k}(r_k - 1) = 2N - 2$$

so

$$\left(1 - \frac{1}{r_1}\right) + \cdots + \left(1 - \frac{1}{r_k}\right) = 2 - \frac{2}{N}$$

We can bound the right-hand side by $[1, 2)$ if $G \neq \{1\}$, and each of the terms on the left-hand side is in $[\frac{1}{2}, 1)$ (since they are each nontrivial). Therefore, we must have either two or three orbits, i.e., $k = 2$ or 3.

Two Orbits

Then,

$$\left(1 - \frac{1}{r_1}\right) + \left(1 - \frac{1}{r_2}\right) = 2 - \frac{2}{N}$$

meaning

$$\frac{1}{r_1} + \frac{1}{r_2} = \frac{2}{N}$$

and since $r_i \leq N$, it must be the case that $r_1 = r_2 = N$, meaning $n_1 = n_2 = 1$. Then, the number of poles is 2, and each pole is fixed by the whole group, so $G = C_n$.

Three Orbits

Then,

$$\left(1-\frac{1}{r_1}\right)+\left(1-\frac{1}{r_2}\right)+\left(1-\frac{1}{r_3}\right)=2-\frac{2}{N}$$

and without loss of generality, $r_1 \le r_2 \le r_3$. Then,

$$\frac{1}{r_1}+\frac{1}{r_2}+\frac{1}{r_3}=1+\frac{2}{N}$$

If $r_1 \ge 3$, then the left-hand side is $\le \frac{1}{3}+\frac{1}{3}+\frac{1}{3}=1$, which is too small. Therefore, $r_1 < 3$, meaning we must have $r_1 = 2$. Thus,

$$\frac{1}{r_2}+\frac{1}{r_3}=\frac{1}{2}+\frac{2}{N}$$

The same argument says that if $r_2 \ge 4$, then the left-hand side is $\le \frac{1}{4}+\frac{1}{4} \le \frac{1}{2}$, which is too small. Therefore, $2 \le r_2 < 4$.

First we consider the subcase where $\underline{r_2 = 2}$:

$$\frac{1}{r_3}=\frac{2}{N}, \text{ so } r_3=\frac{N}{2}$$

and we can determine that $G = D_{N/2}$

Next we consider the subcase where $\underline{r_2 = 3}$:

$$\frac{1}{r_3}=\frac{1}{6}+\frac{2}{N}$$

so $r_3 < 6$, meaning we have subcases where $r_3 = 3, 4$, or 5. We can focus on the case where $r_3 = 4$:

$$\frac{1}{4}=\frac{1}{6}+\frac{2}{N}, \text{ so } N=24.$$

The r_i (stabilizer sizes) are $2, 3, 4$ and the n_i (orbit sizes) are $12, 8, 6$. Suppose we choose a pole p in the size 6 orbit. Then,

$$\text{Stab}(p) = C_4$$

and we can observe that G preserves the set of vertices of an octahedron.

§16 October 9, 2020

Continuing our discussion from last time, we could also consider the finite subgroups of O_3, which is generated by SO_3 and $\{\pm I\}$ with intersection $\{I\}$ since $\det(-I) = (-1)^3 = -1$. Furthermore, $gh = hg$ for all $g \in SO_3$ and $h \in \{\pm I\}$ (commutes). Then,

$$O_3 \simeq SO_3 \times \{\pm I\}$$

allowing us to classify the finite subgroups of O_3 (Goursat's lemma).

Fact

We think of two actions of G (considered as a group) on G (considered as a set):

(1) Left Multiplication: performs as $m_g : x \mapsto gx$. We get a homomorphism from

$$\begin{aligned} G &\to \mathrm{Perm}(G) \\ g &\mapsto m_g \end{aligned}$$

(2) Conjugation: conjugates by g: $x \mapsto gxg^{-1}$. We have

$$\mathrm{orbit}(x) = \{gxg^{-1} : g \in G\} =: C(x)$$

as the **conjugacy class** of x and

$$\mathrm{Stab}(x) = \{g \in G : gxg^{-1} = x\} = \{g \in G : gx = xg\} =: Z(x)$$

as the **centralizer** (in this particular case, it commutes) of x (also a subgroup of G here). Notice that the center $Z \leq Z(x) \leq G$.

(1) If g is in the kernel of our homomorphism, $m_g = \mathrm{id}$, so $m_g(1) = 1$, meaning $g = 1$, so $\ker = \{1\}$, meaning our homomorphism is injective and would be an isomoprhism. This leads to Cayley's theorem.

Theorem 40 (Cayley's Theorem)

If $|G| = n$, then G is isomorphic to a subgroup of S_n.

(2) By the orbit-stabilizer theorem,

$$|G| = |C(x)|\,|Z(x)|$$

In an extreme case, $C(x) = \{x\} \iff Z(x) = G$. We can also consider a class equation based on the set:

Definition (Class Equation)

$$|G| = \underbrace{|C_1| + \cdots + |C_k|}_{\text{conjugacy classes}}$$

where each $|C_i|$ divides $|G|$.

Example

Consider $G = D_4 = \{1, x, x^2, x^3, y, xy, x^2y, x^3y\}$.

Then, 1 and x^2 are both elements of the center, and we have the conjugacy classes

$$\{x, x^3\}, \{y, x^2y\}, \{xy, x^3y\}$$

so for our class equation, we can write

$$8 = \underbrace{1+1}_{\text{center}} + 2 + 2 + 2$$

Definition (p-group)

A finite group G is a **p-group** (for p some prime number) if $|G|$ is a power of p.

One example would be $\begin{pmatrix} 1 & * & * \\ & 1 & * \\ & & 1 \end{pmatrix} \leq GL_3(\mathbb{F}_p)$, which gives a subgroup of order p^3. It is debated whether $\{1\}$ is a p-group; Artin says no, but most people, including Poonen, say yes.

Definition (Elementary Abelian p-group)

An **elementary abelian p-group** is an abelian group G in which every element has order dividing p.

One example would be $C_p \times C_p \times C_p \simeq (\mathbb{F}_p)^3$.

Proposition 41

For G a nontrivial p-group, its center is not just $\{1\}$.

Proof. Given $|G| = p^e$ for some $e \geq 1$, each $|C_i|$ divides p^e, and hence is a power of p (possibly 1). Then, the class equation looks like

$$p^e = \underbrace{(1 + \cdots)}_{|Z|} + \sum(\text{higher powers of } p)$$

Considering the class equation modulo p, it must be the case that p divides $|Z|$ (the order of the *center*). Thus $|Z| \neq 1$ and must be nontrivial (since it contains at least the identity). □

Example

We consider examples of p-group based on power.

- Groups of order p^0: $\{1\}$
- Groups of order p^1 : $C_p \simeq \mathbb{Z}/p\mathbb{Z} \simeq \mathbb{F}_p$ (one possibility)
- Groups of order p^2 : $C_{p^2}, C_p \times C_p$

We show that these are the two distinct groups of order p^2.

Proposition 42

If $|G| = p^2$, then G is abelian.

Proof. The center Z must be strictly larger than $\{1\}$ and is a subgroup of G, i.e., $\{1\} < Z \leq G$, meaning the order of Z is either p or p^2.

If $|Z| = p^2$
Then $G = Z$, which is abelian.

If $|Z| = p$
Choose $x \in G \backslash Z$. Then, $Z < Z(x) \leq G$, since $x \in Z(x)$, so we must have $Z(x) = G$, meaning x commutes with everything, meaning $x \in Z$, which is a contradiction. □

Corollary

If $|G| = p^2$, then $G \simeq C_{p^2}$ or $C_p \times C_p$.

Proof. Every element of G has order dividing p^2. We consider the following cases

There exists $a \in G$ of order p^2
Then $\langle a \rangle$ (group generated by a) has order p^2, so $\langle a \rangle = G$, so G is cyclic.

Every element has order dividing p
Then G is an elementary abelian p-group, so G is a 2-dimensional $\mathbb{F}_p$-vector space, so $G \simeq (\mathbb{F}_p)^2 \simeq C_p \times C_p$. □

Fact

If d is a divisor of n, then n factors into $d, \frac{n}{d}$. If N is a normal subgroup of G, then "G decomposes into N, G/N."

Definition (Simple Group)

If n is a *prime number*, $n > 1$ and its only positive divisors are $1, n$. Analogously, G is a **simple group** if and only if $G \neq \{1\}$ and its only normal subgroups are $\{1\}, G$.

Consider C_{10} where C_5 is a normal subgroup, so C_{10} is not simple. In general, C_n is simple if and only if n is prime.

§17 October 13, 2020

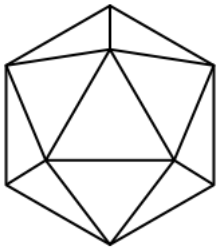
Theorem 43 (Icosahedral Group)

The icosahedral group $I \simeq A_5$, and it is simple.

Proof. Recall

$$I = \{\text{rot. symmetries of an icosahedron}\} \leq SO_3$$

The three types of rotations we consider of an icosahedron have poles that go through the midpoints of opposite edges, opposite vertices, or the centers of faces:



Pole direction	|Orbit|	Stabilizer	Elements $\neq 1$ in stabilizer	How many elts.?
midpt. of edge	30	C_2	$\rho_{(u,\pi)}$	15
center of face	20	C_3	$\rho_{(u,\frac{2\pi}{3})}, \rho_{(u,\frac{4\pi}{3})}$	20
vertex	12	C_5	$\rho_{(u,\frac{2\pi}{5})}, \ldots, \rho_{(u,\frac{8\pi}{5})}$	$2 \cdot 12$

where we have omitted the identity. Then, the class equation yields

$$60 = 1 + 12 + 12 + 15 + 20$$

We can argue that this is the class equation as follows: If u, u' are poles in the same orbit, $\rho_{(u,\pi)}$ is conjugate to $\rho_{(u',\pi)}$. More generally, if $u' = gu$, then $\rho_{(u',\pi)} = g\rho_{(u,\pi)}g^{-1}$ and similarly for all the other angles. Thus each of the number of elements corresponds to one class because the orbits were grouped according to the same orbit.

If $N \lhd I$, then N is the union of some of the conjugacy classes but also $|N|$ divides the order of the group, meaning in this case, either $N = \{1\}$ or I, so I is simple.

Also, I acts on a dodecahedron, and we can let $S :=$ the set of 5 cubes with vertices at the dodecahedron's vertices. Then I acts on S with a homomorphism:

$$\phi : I \to \text{Perm}(S) \simeq S_5$$

Since a kernel is a normal subgroup, $\ker\phi \lhd I$, and since I is simple, either $\ker\phi = \{1\}$ or $\ker\phi = I$. The latter is impossible since a non-trivial rotation that moves some cube can easily be found in I. Then, $\ker\phi = \{1\}$ so $\phi : I \to S_5$ is injective. Consider

$$I \xrightarrow{\phi} S_5 \xrightarrow{\text{sign}} \{\pm 1\}$$

where $I \xrightarrow{\psi} \{\pm 1\}$ meaning $\ker\psi = \{1\}$ or $\ker\psi = I$. It is impossible for ψ to be injective since I has 60 elements and $\{\pm 1\}$ has 2. Therefore, $\ker\psi = I$. Then,

$$\text{sign}(\phi(x)) = 1$$

so $\phi(x) \in A_5$ and $I \hookrightarrow A_5$, meaning $I \simeq A_5$ since both have order 60. □

Example (Conjugation in S_n)

Consider

$$p = (14)(253) \in S_5 = \text{Perm}(\{1,2,3,4,5\}) \simeq \text{Perm}(\{a,b,c,d,e\})$$

and

$$p' = (ad)(bec) \in \text{Perm}(\{a,b,c,d,e\})$$

We can define a bijection

$$q : \{1,2,3,4,5\} \to \{a,b,c,d,e\}$$

mapping each element to the one in its corresponding position. Then,

$$qpq^{-1} = p'$$

Similarly, if q is the bijection $\{1,2,3,4,5\} \to \{1,2,3,4,5\}$ where $1 \mapsto 4$, $2 \mapsto 1$, $3 \mapsto 3$, $4 \mapsto 2$, $5 \mapsto 5$ and

$$p = (14)(253)$$

then

$$qpq^{-1} = (42)(153)$$

Fact

This yields the conclusion

- p is conjugate to $p' \iff$ the lengths of the cycles for p are the same as for p'
- Each conjugacy class in S_n is the set of permutations with a given collection of cycle lengths

Example

How many elements are in the conjugacy class of $x := (1324)$ in S_4?

The conjugacy class of x is the set of all 4-cycles in S_4. One way we can determine the number of elements in $C(x)$ is with the orbit-stabilizer theorem:

$$|S_4| = |C(x)|\,|Z(x)|$$

where

$$Z(x) = \{p \in S_4 : pxp^{-1} = x\} = \{p \in S_4 : \text{relabelling } x \text{ using } p \text{ gives } x\}$$

and relabelling x to give x again is just the number of ways to pick a starting point, so $|Z(x)| = 4$, meaning

$$|C(x)| = 4!/4 = 6.$$

Example

How many elements of S_{10} have the same cycle type as $(123)(456)(789\ 10)$?

We have

$$|Z(x)| = 2! \cdot 3 \cdot 3 \cdot 4$$

where 2! is the number of ways to decide which 3-cycle is first, and the remaining elements indicate the number of ways to choose a starting point every cycle. Then,

$$|C(x)| = \frac{|S_{10}|}{|Z(x)|} = \frac{10!}{2! \cdot 3 \cdot 3 \cdot 4} = 50400.$$

§18 October 14, 2020

We first give a quick review (to help on the problem set!):

Fact

Every element of O_2 has the form ρr where

- ρ is a rotation around 0, and
- $r = 1$ or $r =$ a reflection in the x-axis.

Fact

Every element of M_2 has the form $t_a \rho r$ where

- t_a is a translation,
- ρ is a rotation around 0, and
- $r = 1$ or $r =$ a reflection in the x-axis.

To conjugate x by $t_a \rho r$, conjugate x by r, then by ρ, and then by t_a.

Example

S_4 acts by conjugation on the conjugacy class of $(12)(34)$:

$$\{a = (12)(34), b = (13)(24), c = (14)(23)\}$$

We can consider the permutation representation:

$$\phi : S_4 \to \mathrm{Perm}(\{a, b, c\}) \simeq S_3$$

where, as some examples of S_4 acting on this set,

$$(12) \mapsto (bc), (123) \mapsto (acb), (12)(34) \mapsto 1$$

The image in this case must generate all of S_3, so ϕ is surjective. We can find that

$$|\ker \phi| = \frac{|S_4|}{|\mathrm{im}\, \phi|} = \frac{4!}{3!} = 4.$$

In fact,

$$V := \ker \phi = \{1, a, b, c\}$$

is a normal subgroup of S_4 with $|V| = p^2$ with $p = 2$, so either $V \simeq C_4$ or $C_2 \times C_2$, and in this case $V \simeq C_2 \times C_2$.

Example
Is A_4 simple?

A_4 is not simple because $V \lhd A_4$.

Example
Consider the conjugacy classes in A_4.

We could consider the conjugacy classes

$$\{1\}, \{3-\text{cycles}\}, \{a, b, c\}$$

but this would give a class equation

$$12 = 1 + 8 + 3$$

and 8 does not divide 12. In particular, the issue arises from 3-cycles. (123) is conjugate to (213) in S_4, by conjugation by (12), but they are not conjugate in A_4 since (12) is not in A_4. Thus, the 3-cycles break down into two conjugate classes. Notice that they *would* be conjugates in A_5 via conjugation by (12)(45).

Lemma
In particular, for $n \geq 5$, {3-cycles in A_n} form one conjugacy class in A_n.

Theorem 44 ($n \geq 5 \implies A_n$ simple)
If $n \geq 5$, then A_n is simple.

Proof. Suppose $\{1\} \neq N \lhd A_n$. We show that $N = A_n$. We can choose $x \in N$, with $x \neq 1$ such that it fixes the most elements in $\{1, 2, \ldots, n\}$.

Case: x is a 3-cycle
Then N contains all conjugates of x, i.e., all the 3-cycles, which generate A_n (proved on Problem Set 1), so $N = A_n$.

Case: x is a product of disjoint 2-cycles
Without loss of generality, $x = (12)(34)$[maybe more]. Let $g = (345) \in A_n$. We can form

$$y := \underbrace{gxg^{-1}}_{\text{in } N}\underbrace{x^{-1}}_{\text{in } N}.$$

Then $y \in N$ and $y(3) = 5$, so $y \neq 1$. Also, y fixes 1, 2, and everything > 5 fixed by x, so y fixes more elements than x does, leading to a contradiction.

Case: x is $(123\dots)$ and moves 4 and 5

Let $g = (345) \in A_n$ and form

$$y := gxg^{-1}x^{-1} \in N.$$

Then $y(3) = 4$, so $y \neq 1$ and y fixes 2 and everything fixed by x, so y fixes more elements than x, leading to a contradiction. □

Fact

If G acts on G by conjugation, G also acts on $\{\text{subsets of } G\}$ by conjugation, and also acts on $\{\text{subgroups of } G\}$ by conjugation.

If H is a subgroup, then $gHg^{-1} = \text{inn}_g(H)$, the image of H under a homomorphism, which is a subgroup too.

Definition (Normalizer)

For $H \leq G$, we consider the stabilizer for the conjugation action defined above:

$$\text{Stab}(H) = N(H) = \{g \in G : \underbrace{gHg^{-1} = H}_{g \text{ normalizes } H}\}$$

and the whole set is called the **normalizer** $N(H)$. Then, $H \lhd N(H)$.

Example

Consider $G = S_3$.

The subgroup $\langle(123)\rangle$ is normal while $\langle(12)\rangle, \langle(13)\rangle, \langle(23)\rangle$ are conjugate subgroups. Consider $H = \langle(12)\rangle$. By the orbit-stabilizer theorem,

$$|G| = (\text{number of conjugates of } H)\,|N(H)|$$

or

$$6 = 3\,|N(H)|$$

so $|N(H)| = 2$, meaning $N(H) = H$.

Fact

$$H \text{ is normal in } G \iff gHg^{-1} = H \text{ for all } g \in G \iff N(H) = G.$$

Theorem (Lagrange's Theorem)
If $H \leq G$ a finite group, then $|H|$ is a divisor of $|G|$.

However, if d divides $|G|$, G need not have a subgroup of order d. As an example, $G = A_4$ has order 12, but G has no subgroup of order 6 (Problem Set 7). However, this does hold in a special case.

§19 October 16, 2020

Theorem 45 (First Sylow Theorem)
Suppose $|G| = n = p^e m$ where p^e is the largest power of p dividing n, and m is not divisible by p. Then, G has a subgroup of order p^e, i.e., a **Sylow p-subgroup** of G.

Proof. We utilize our proof of Cauchy's Theorem for the special case (Theorem 46). We proceed by induction. Assume that p divides $|G|$ (otherwise $\{1\}$ is a Sylow p-subgroup). We can consider the class equation:

$$|G| = \underbrace{1 + 1 + \cdots + 1}_{|Z|} + |C(x)| + \cdots + |C(z)|$$

Case 1: Some $|C(x)| > 1$ is not divisible by p
The equation

$$|G| = |C(x)|\,|Z(x)|$$

shows that $|Z(x)|$ is divisble by the same power of p as $|G|$, so by our inductive hypothesis, there exists a Sylow p-subgroup of $Z(x)$, which must be a Sylow p-subgroup of G.

Case 2: Every $|C(x)| > 1$ is divisible by p
The class equation says that $|Z|$ is divisible by p, and Cauchy's Theorem for $|Z|$ (abelian and finite) says that there exists $C_p \leq Z$. Then, we get

$$G \xrightarrow{\pi} G/C_p$$

and there exists (by the inductive hypothesis)

$$S := \text{ a Sylow } p\text{-subgroup of } G/C_p$$

so $|S| = p^{e-1}$. Then,

$$\left|\pi^{-1}(S)\right| = |\ker \pi|\,|S| = p \cdot p^{e-1} = p^e$$

so $\pi^{-1}(S)$ is a Sylow p-subgroup of G.

□

Example

Consider $G = S_4$ where $|G| = 24 = 2^3 \cdot 3$.

We can consider

$$\langle (12), (34), (13)(24) \rangle$$

which is a Sylow 2-subgroup of S_4.

Theorem 46 (Cauchy's Theorem in Group Theory)

If prime p divides $|G|$, then G has an element of order p. Equivalently, G has a subgroup of order p.

Proof. We first prove Cauchy's Theorem for finite abelian G.

Let $x_1, \ldots, x_n$ be the elements of G. Then we have a surjective homomorphism

$$\begin{aligned} \langle x_1 \rangle \times \langle x_2 \rangle \times \cdots \times \langle x_n \rangle &\to G \\ (y_1, y_2, \ldots, y_n) &\mapsto y_1 + \cdot + y_n \end{aligned}$$

where y_i are the multiples generated. If p divides $|G|$, then p divides $|\langle x_1 \rangle \times \langle x_2 \rangle \times \cdots \times \langle x_n \rangle|$, so p divides some $|\langle x_i \rangle|$. Then, $\langle x_i \rangle$ contains a copy of C_p. □

Proof. We now prove Cauchy's Theorem for every finite group G.

Given p divides $|G|$ we can choose

$$S := \text{ a Sylow } p\text{-subgroup of } G$$

of order $p^e > 1$ (based on the First Sylow Theorem). Within S, we choose a cyclic a cyclic subgroup

$$\langle x \rangle \text{ where } x \in S - \{1\}$$

so $1 \neq \text{ord}(x) \mid p^e$, so $\text{ord}(x) = p^a$, where $0 < a \leq e$. Then, as a subgroup of $\langle x \rangle$, we can find C_p, a subgroup of order p. □

Theorem 47 (Second Sylow Theorem)

Fixing a finite group G and a prime p,

(a) The Sylow p-subgroups are conjugates of each other

(b) Every p-subgroup of G is contained in a Sylow p-subgroup

(c) The Sylow p-subgroups are normal if and only if G has a unique Sylow p-subgroup

Proof. We begin with a lemma.

Lemma (Fixed Point Lemma)

Suppose X is a finite set and P is a p-group acting on X. Then

$$X^P := \{x \in X \text{ fixed by } P\}$$

Then

$$|X| \equiv \left|X^P\right| \pmod{p}$$

Proof. Every orbit in $X - X^P$ has size p^a for some $a \geq 1$. Adding their sizes yields that $\left|X - X^P\right|$ is divisible by p. □

(b) Let S be a Sylow p-subgroup and let P be any p-subgroup. Let $X = G/S$, so

$$\left|X^P\right| = |X| = m \not\equiv 0 \pmod{p}$$

so there exists $x \in X$ fixed by P (nonzero order). Then,

$$P \leq \text{Stab}(x) = \underbrace{aSa^{-1}}_{\text{another Sylow}} \quad \text{for some } a \in G$$

(a) Then if S' is another Sylow p-subgroup, the proof of (b) says $S' \leq aSa^{-1}$ for some a (and equality must be satisfied since both have order p^e). Thus, all Sylow p-subgroups will be conjugate to each other.

(c) S is normal if and only if $gSg^{-1} = S$ for all $g \in G$. This is equivalent to saying that every Sylow p-subgroup equals S, implying that G has only one Sylow p-subgroup. □

Theorem 48 (Third Sylow Theorem)

The number of Sylow p-subgroups is 1 mod p and divides m.

Proof. Let $X = \{\text{Sylow } p\text{-subgroups}\}$ and choose $S \in X$. Then, G acts by conjugation on X. Then,

$$|G| = |\text{orbit}(S)|\,|\text{Stab}(S)| = |X|\,|N(S)|$$

This means that

$$|X| = \frac{|G|}{|N(S)|} \mid \frac{|G|}{|S|} = \frac{p^e m}{p^e} = m$$

since $S \leq N$, so $|S| \mid |N(S)|$. Next consider restricting the conjugation of G on X to get an action of S on X. For $S' \in X$,

$$S' \in X^S \iff S' \text{ is fixed by every element of } S$$

which is equivalent to saying every element of S is in $\mathrm{Stab}(S')$, meaning $S \leq N(S')$. We claim then that $S' = S$. In particular, $S' \lhd N(S')$ and is a Sylow of $N(S')$. Then, by the Second Sylow Theorem, S' is the *only* Sylow of $N(S')$, so $S = S'$. Thus,

$$X^S = \{S\}$$

and by the Fixed Point Lemma,

$$|X| \equiv \left|X^S\right| \pmod{p}$$

so $|X| \equiv 1 \pmod{p}$. □

§20 October 19, 2020

Theorem 49

Suppose G is a finite abelian group. Then G is isomorphic to the direct product of its Sylow p-subgroups, one for each prime p dividing $|G|$.

Proof. First, write

$$|G| = p_1^{e_1} \cdots p_k^{e_k}$$

and let S_i be *the* (automatically normal because in an abelian group, and hence unique by Second Sylow Theorem) Sylow p_i-subgroup. Since G is abelian, consider the homomorphism

$$\begin{aligned} \phi : S_1 \times \cdots \times S_k &\to G \\ (x_1, \ldots, x_k) &\mapsto x_1 + \cdots + x_k \end{aligned}$$

Then $\operatorname{im}\phi$ contains each S_i so $|\operatorname{im}\phi|$ is divisible by $p_i^{e_i}$ for all i, and hence is divisible by $|G|$ (the product of all prime powers), meaning $\operatorname{im}\phi = G$, so ϕ is surjective. Furthermore,

$$|S_1 \times \cdots \times S_k| = p_1^{e_1} \cdots p_k^{e_k} = |G|,$$

so ϕ must be bijective, meaning ϕ is an isomorphism. □

Example

Consider $G = C_6 \times C_4 = \underbrace{C_2 \times C_4}_{\text{Sylow 2-subgp.}} \times \underbrace{C_3}_{\text{Sylow 3-subgp.}}$

Example

$\mathbb{F}_7^\times$ is an abelian group of order 6, so

$$\mathbb{F}_7^\times \simeq C_2 \times C_3 \simeq C_6$$

and hence $\mathbb{F}_7^\times$ contains a copy of C_3: $\{1, 2, 4\}$.

Example (The $ax + b$ Group)

Consider

$$M := \{ax + b \mid a \in \mathbb{F}_7^\times \text{ and } b \in \mathbb{F}_7\} \leq \text{Perm}(\mathbb{F}_7)$$

which is a group under composition.

For instance,

$$(5x + 2) \circ (3x + 1) = \text{id}$$

since $5(3x + 1) + 2 = 15x + 7 = x$. In particular,

$$|M| = (7 - 1) \cdot 7 = 42.$$

There is a homomorphism

$$\begin{aligned} M &\xrightarrow{\pi} \mathbb{F}_7^\times \\ (x \mapsto ax + b) &\mapsto a \end{aligned}$$

and has kernel

$$L := \{(x \mapsto x + b) : b \in \mathbb{F}_7\} \simeq \mathbb{F}_7$$

Let $J := \pi^{-1}(\{1, 2, 4\}) = \{(ax + b) \in M : a \in \{1, 2, 4\}, b \in \mathbb{F}_7\}$. Then,

- $|J| = 21$
- J is nonabelian
- L is the Sylow 7-subgroup of J (kernel is normal and hence unique Sylow)
- The 14 elements of $J \backslash L$ have order 3 (can't have order 1, 7, or 21)
- There are 7 subgroups of order 3

Theorem 50 (Groups of Order 21)

Every group of order 21 is isomorphic to C_{21} or J (as defined above).

Proof. Let G be a group of order $21 = 3 \cdot 7$. By the Third Sylow Theorem, the number of Sylow 7-subgroups is 1 mod 7, but also divides 3, so it must be 1. Let $H_7 = \langle x \rangle$ be the Sylow 7-subgroup. By the Second Sylow Theorem, H_7 is normal in G.

Similarly, the number of Sylow 3-subgroups is 1 mod 3 and divides 7, so it could be 1 or 7. Let $H_3 = \langle y \rangle$ be a Sylow 3-subgroup. Then,

$$H_7 \cap H_3 = \{1\} \text{ (order divides 7 and 3)}$$

and hence

$$H_7 \times H_3 \to G \text{ is injective (but not necessarily a homomorphism)}$$

and also bijective (because there are 21 elements on both sides). Thus,

$$G = \{x^i y^j : 0 \leq i < 7, 0 \leq j < 3\} \text{ as a set.}$$

Since H_7 is normal,

$$yxy^{-1} = x^a \text{ for some } a \in \{1, \ldots, 6\} = \mathbb{F}_7^\times.$$

We claim that for each a, there is at most one possible G up to isomorphism. In particular, given a,

$$yx = x^a y$$

rewrites any product $(x^i y^j)(x^{i'} y^{j'})$ as $x^c y^d$, so the group law is determined. In fact, we can get a homomorphism

$$\phi : H_3 \to \mathrm{Aut}(H_7) \simeq \mathbb{F}_7^\times$$
$$Y \mapsto (\text{conjugation by } Y) \leftrightarrow (b \text{ such that } YxY^{-1} = x^b)$$

Case 1: $|\mathrm{im}\,\phi| = 1$
Then $\phi(y) = 1$, so $yxy^{-1} = x^1$, meaning $yx = xy$, so $G \simeq \langle x \rangle \times \langle y \rangle \simeq C_7 \times C_3 \simeq C_{21}$.

Case 2: $|\mathrm{im}\,\phi| = 3$
Then ϕ maps $\langle y \rangle$ to $\{1, 2, 4\} \leq \mathbb{F}_7^\times$ (the only subgroup of order 3). Choosing a different generator (y) of H_3 if necessary, without loss of generality, $\phi(y) = 2$ (it either maps to 2 or 4). Then, $yxy^{-1} = x^2$. We know that such a group exists, because J exists. □

Corollary

The same argument shows that given primes $p < q$, the groups of order pq are

- C_{pq}
- If $p \mid q - 1$, also $\{ax + b : a \in \mathbb{F}_q^\times (\mathrm{ord}(a) = p), b \in \mathbb{F}_q\}$

§21 October 21, 2020

Definition (Free Groups)

A **free group** has elements with no constraints or relations. We start with certain symbols, and a **word** is a finite string of these symbols, with repetition allowed.

For example, the free group on 2 generators is

$$F_2 := \langle x, y \mid \rangle$$

We start with symbols x, y, x^{-1}, y^{-1} (just symbols, not necessarily inverses). Some words we could make are:

- $xyx^{-1}y^{-1}$
- $xxyyy$
- $xxyyyyy^{-1}$
- 1 (the empty word)
- $x^{-1}yy^{-1}x$

Sometimes, different words should represent the same element of the group we are constructing.

Definition (Reduction)

A word is **reduced** if it doesn't have an adjacent pair of symbols that have the potential to cancel.

Fact

We consider the properties of reducing words:

- Given any word w, one can repeatedly cancel (possibly in more than one way) until reaching a reduced word.
- All ways of cancelling (reducing) lead to the same reduced word.

We prove the second property:

Proof. If w is reduced, then we are done. If not, without loss of generality, there is some

$$\cdots \underline{xx^{-1}} \cdots$$

We claim that every reduced form can be reached by cancelling xx^{-1} first. Start with some cancellation sequence. If $\underline{xx^{-1}}$ gets cancelled at some point, we could have cancelled it first and reached the same reduced form.

If, instead, $\underline{xx^{-1}}$ never gets cancelled, then some stage in the sequence must contain either

$$x^{-1}\underline{xx^{-1}} \text{ or } \underline{xx^{-1}}x$$

and one of the $x^{-1}x$ is cancelled, but at that stage, one could also cancel $\underline{xx^{-1}}$ and obtain the same result, allowing us to ultimately obtain the same reduced form.

Therefore, we might as well cancel xx^{-1} first and proceed. Similarly, we can at any point cancel any terms and obtain the same possible reduced words, meaning in the end, there is only one possible final reduced word (we could prove this formally with induction).

Example

Then, we can better express

$$F_2 := \{\text{words}\}/\sim$$

as the set of equivalence classes where $w \sim w'$ means that w and w' have the same reduced form. Multiplication in F_2 comes from concatenation of words.

Concatenation of words is well-defined in F_2 because if $w \sim w'$ and $v \sim v'$, then

$$wv \sim w'v'.$$

We can reduce wv by first cancelling as much as possible within w and within v; similarly, we can reduce $w'v'$ by first cancelling as much as possible within w' and within v'. This stage obtained from both will be the same, and we can then reduce them to the same word.

This group is associative, because concatenation is associative; has an identity, namely the equivalence class of the empty word; and has inverses (easy to generate a cancelling word given any word). □

§22 October 23, 2020

Fact

F_2 is generated by $[x]$ and $[y]$ (their equivalence classes), which satisfy no relations, except those forced by the group axioms.

Proposition 51 (Universal Mapping Property of Free Group)

Let G be a group. Giving a homomorphism $\phi : F_2 \to G$ is the same as giving $g, h \in G$ (in order, possibly equal).

Proof. Given ϕ, let

$$g := \phi([x])$$
$$h := \phi([y])$$

On the other hand, given $g, h \in G$, define $\phi : F_2 \to G$ by sending

$$[xyyx^{-1}] \mapsto ghhg^{-1}$$

where the first element is a word, and the second is evaluated in G. □

Fact

In many cases, we have a set of **relations** R:

$$\langle x, y \mid R \rangle$$

where each element of R is forced to be 1 (we can think of them as elements of the free group).

For example, we can consider the symmetric group

$$\langle x, y \mid x^3 = 1, y^2 = 1, yx = x^2 y \rangle$$

or equivalently

$$R := \{x^3, y^2, yxy^{-1}x^2\} \subset F_2.$$

We can define $\mathcal{R}$ as the normal subgroup generated by R, i.e., the set of elements obtained from R by repeatedly taking products, inverses, and conjugating by elements of F_2, which is in fact the smallest normal subgroup containing R. $\mathcal{R}$

Definition

Then, we can define

$$\langle x, y \mid R \rangle := F_2/\mathcal{R}.$$

Proposition 52 (Universal Mapping Property of $\langle x, y \mid R \rangle$)

Let G be a group. Giving a homomorphism $\phi : \langle x, y \mid R \rangle \to G$ is the same as giving $g, h \in G$ satisfying each relation in R.

Proof. The following are equivalent:

- $\phi : \langle x, y \mid R \rangle \to G$
- $F_2/\mathcal{R} \to G$
- $F_2 \to G$ mapping $\mathcal{R}$ to 1 (universal property of quotient group)
- $F_2 \to G$ mapping R to 1
- $g, h \in G$ satisfying each relation in R (universal property of free group)

□

Corollary

The homomorphism ϕ is surjective if and only if g, h generate G.

Example

Consider

$$\langle x, y \mid x^5 = 1, y^2 = 1, yxy^{-1} = x^{-1} \rangle \xrightarrow{\phi} D_5$$

We have

$$\begin{aligned} x &\mapsto \text{ rotation by } \frac{2\pi}{5} \\ y &\mapsto \text{ reflection in horiz. axis} \end{aligned}$$

Then, ϕ is

- well-defined, since the rotation and reflection satisfy the relations
- surjective, since rotations and reflections generate D_5
- injective, since using relations allows us to write every element as $x^i y^j$ where $0 \leq i < 5$ and $0 \leq j < 2$, each of which is distinct in D_5

Thus ϕ is an isomorphism.

Example

Consider the group

$$\langle x, y \mid xy = yx \rangle \simeq \mathbb{Z}^2$$

which is a *free abelian group*.

Elements in the group can be written as $x^i y^j$ with $i, j \in \mathbb{Z}$. We can map

$$x \mapsto (1, 0)$$

$$y \mapsto (0, 1)$$

where elements in $\mathbb{Z}^2$ are combined additively.

Example

Consider

$$\langle x, y \mid x^7 = 1, y^3 = 1, yxy^{-1} = x^3 \rangle$$

In this group,

$$x = y^3 x y^{-3} = y(y(yxy^{-1})y^{-1})y^{-1} = ((x^3)^3)^3 = x^{27}$$

so $x^{26} = 1$, meaning $\text{ord}(x) \mid 26$ but also $\text{ord}(x) \mid 7$, so $\text{ord}(x) = 1$, meaning $x = 1$. Then, the group is equivalent to

$$\langle y \mid y^3 = 1 \rangle \simeq C_3.$$

Next, we consider bilinear forms and first provide a classic example, the dot product:

Example

The dot product maps

$$\begin{aligned} \mathbb{R}^2 \times \mathbb{R}^2 &\xrightarrow{\cdot} \mathbb{R} \\ x, y &\mapsto x_1 y_1 + x_2 y_2 \end{aligned}$$

If (y_1, y_2) is specialized to some value, say $(7, 8)$, this becomes a linear function

$$\begin{aligned} \mathbb{R}^2 &\to \mathbb{R} \\ (x_1, x_2) &\mapsto 7x_1 + 8x_2 \end{aligned}$$

and vice-versa (if x is specialized to some value).

Definition

A **bilinear form** (*bilinear pairing*) on V ($\mathbb{R}$-vector space) is a function

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R}$$

that is linear in the first variable, i.e., for all $v_1, v_2, w \in V$,

$$\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle$$

and for all $v, w \in V$ and $r \in \mathbb{R}$,

$$\langle rv, w\rangle = r\langle v, w\rangle$$

and similarly, linear in the second variable (symmetric with operations in the second variable).

Fact

Recall the bijective correspondence

$$\{m \times n \text{ matrices}\} \leftrightarrow \{\text{linear transformations } \mathbb{R}^n \times \mathbb{R}^m\}$$

Proposition 53

There exists a bijective correspondence

$$\{n \times n \text{ matrices}\} \leftrightarrow \{\text{bilinear forms on } \mathbb{R}^n\}$$

where given A (an $n \times n$ matrix), $\langle x, y\rangle := x^t A y$, and given $\langle\ ,\ \rangle$, define $A = (a_{ij})$ with $a_{ij} = \langle e_i, e_j\rangle$ (using standard basis vectors of $\mathbb{R}^n$).

§23 October 26, 2020

Proposition 54

More generally, given $\langle\ ,\ \rangle_V$ on V and a basis $B = (v_1, \ldots, v_n)$ of V, we get the matrix of $\langle\ ,\ \rangle_V$ with respect to B.

Method 1:

We take $A = (a_{ij})$, where $a_{ij} := \langle v_i, v_j\rangle_V$.

Method 2:

The basis B gives an isomorphism from $\mathbb{R}^n \to V$, meaning there exists $\langle\ ,\ \rangle$ in $\mathbb{R}^n$ which corresponds to $\langle\ ,\ \rangle_V$ in V. Then, we define A so that

$$x^t Ay = \langle x, y\rangle := \langle Bx, By\rangle_V.$$

Example (Change of Basis)

Suppose B' is another basis of V where $B' = BP$. If the matrix of the bilinear form with respect to B is A, the matrix of the bilinear form with respect to B' is

$$P^t AP$$

We can coordinatize V in two ways from $\mathbb{R}^n$ (using B or B'), and there is a basechange matrix $P \in GL_n(\mathbb{R})$ relating $\mathbb{R}^n$ to $\mathbb{R}^n$. We can let $\langle\ ,\ \rangle$ to be the pairing that is related by B, and $\langle\ ,\ \rangle'$ to be the pairing that is related by B'. Then,

$$\langle x, y\rangle' = \langle Px, Py\rangle = (Px)^t A(Py) = x^t(P^t AP)y.$$

So the matrix of $\langle\ ,\ \rangle_V$ with respect to B' is $P^t AP$.

Definition (Symmetric Bilinear Pairings)

A bilinear pairing $\langle\ ,\ \rangle$ is **symmetric** if

$$\langle v, w\rangle = \langle w, v\rangle$$

for all $v, w \in V$. A matrix $A = (a_{ij})$ is symmetric if

$$a_{ij} = a_{ji}$$

for all i, j, i.e., $A^t = A$. A pairing is symmetric if and only if its matrix is.

Definition (Skew-Symmetric Bilinear Pairings)

A bilinear pairing $\langle\ ,\ \rangle$ is **skew-symmetric** if

$$\langle v, w\rangle = -\langle w, v\rangle$$

for all $v, w \in V$. A matrix A is skew-symmetric if

$$a_{ij} = -a_{ji}$$

for all i, j, i.e., $A^t = -A$. A pairing is skew-symmetric if and only if its matrix is.

Definition (Positive (Semi)Definite)

A symmetric bilinear form $\langle\ ,\ \rangle$ on V is **positive definite** if

$$\langle v, v\rangle > 0$$

for all $v \neq 0$ (e.g. the dot product). The bilinear form is **positive semidefinite** if

$$\langle v, v\rangle \geq 0$$

for all $v \neq 0$.

Definition (Negative (Semi)Definite)

A symmetric bilinear form $\langle\ ,\ \rangle$ on V is **negative definite** if

$$\langle v, v\rangle < 0$$

for all $v \neq 0$ (e.g. the dot product). The bilinear form is **negative semidefinite** if

$$\langle v, v\rangle \leq 0$$

for all $v \neq 0$.

Example (Lorentz Form)

An important example in special relativity is

$$V = \text{ spacetime } = \mathbb{R}^3 \times \mathbb{R} = \mathbb{R}^4$$

which has the Lorentz form:

$$\langle (x_1, y_1, z_1, t_1), (x_2, y_2, z_2, t_2)\rangle := x_1x_2 + y_1y_2 + z_1z_2 - t_1t_2.$$

Since the bilinear form is neither positive semidefinite nor negative semidefinite, it is **indefinite**.

We can consider the following analogs between $\mathbb{R}$ and $\mathbb{C}$:

$\mathbb{R}$	$\mathbb{C}$
$\lvert x\rvert^2 = x^2$	$\lvert z\rvert^2 = \bar{z}z$
$\mathbb{R}^n$	$\mathbb{C}^n$
$A \in \mathbb{R}^{m\times n}$	$A \in \mathbb{C}^{m\times n}$
transpose A^t	adjoint $A^* :=$ complex conjugate of A^t
symmetric matrix: $A^t = A$	Hermitian (self-adjoint) matrix: $A^* = A$
$x \cdot y = x^t y = x_1y_1 + \cdots + x_ny_n$	$\langle x, y\rangle := x^*y = \bar{x_1}y_1 + \cdots + \bar{x_n}y_n$

Example

Consider the analog between a symmetric bilinear form on V and a Hermitian form on V.

For the symmetric bilinear form on V, we have $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ such that for all $c \in \mathbb{R}$,

$$\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle$$

$$\langle cv, w\rangle = c\langle v, w\rangle$$

$$\langle v, w_1 + w_2\rangle = \langle v, w_1\rangle + \langle v, w_2\rangle$$

$$\langle v, cw\rangle = c\langle v, w\rangle$$

$$\langle w, v\rangle = \langle v, w\rangle$$

For the Hermitian form on V, we have $\langle\ ,\ \rangle : V \times V \to \mathbb{C}$ such that for all $c \in \mathbb{C}$,

$$\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle$$

$$\langle cv, w\rangle = \bar{c}\langle v, w\rangle$$

$$\langle v, w_1 + w_2\rangle = \langle v, w_1\rangle + \langle v, w_2\rangle$$

$$\langle v, cw\rangle = c\langle v, w\rangle$$

$$\langle w, v\rangle = \overline{\langle v, w\rangle}$$

where the Hermitian form is conjugate linear in the first variable and $\mathbb{C}$-linear in the second variable. Notice that our last equation implies $\langle v, v\rangle \in \mathbb{R}$. As a result, we can define that $\langle\ ,\ \rangle$ is *positive definite* (for a symmetric or Hermitian form) if

$$\langle v, v\rangle > 0$$

for all $v \neq 0$, and define the other definitions similarly.

Fact

Every symmetric bilinear form on $\mathbb{R}^n$ is, for some symmetric $A \in \mathbb{R}^{n\times n}$,

$$x, y \mapsto x^t A y.$$

Every Hermitian form on $\mathbb{C}^n$ is, for some Hermitian $A \in \mathbb{C}^{n\times n}$,

$$x, y \mapsto x^* A y.$$

Proposition 55

Given $\langle\ ,\ \rangle$ on V with a $\mathbb{C}$-basis $B = (v_1, \ldots, v_n)$ of V, we get a matrix

$$A = (a_{ij}) \text{ where } a_{ij} := \langle v_i, v_j \rangle.$$

Example (Change of Basis)

Suppose B' is another basis of V where $B' = BP$. If the matrix of the bilinear form with respect to B is A, the matrix of the bilinear form with respect to B' is

$$P^* A P$$

§24 October 28, 2020

Definition (Real Orthogonal Matrix)

For $P \in \mathrm{GL}_n(\mathbb{R})$, the following are equivalent:

1. $(Px) \cdot (Py) = x \cdot y$
2. The columns of P form an orthonormal basis of $\mathbb{R}^n$
3. $P^t P = I$
4. $P^t = P^{-1}$

If these hold, P is an **orthogonal matrix**.

Definition

The **orthogonal group** is

$$O_n := \{P \in GL_n(\mathbb{R}) : P^t P = I\}.$$

Definition (Unitary Matrix)

For $P \in GL_n(\mathbb{C})$, the following are equivalent:

1. $\langle Px, Py\rangle = \langle x, y\rangle$ for all $x, y \in \mathbb{C}^n$
2. The columns of P form an orthonormal basis of $\mathbb{C}^n$
3. $P^*P = I$
4. $P^* = P^{-1}$

If these hold, P is a **unitary matrix**.

Definition

The **unitary group** is

$$U_n := \{P \in GL_n(\mathbb{C}) : P^*P = I\}.$$

Theorem 56

For $A \in C^{n\times n}$ a Hermitian matrix, then the eigenvalues of A are real numbers.

Proof. Suppose that

$$Av = \lambda v$$

for eigenvalue λ and eigenvector v in $\mathbb{C}^n$ (not 0). Taking the adjoints yields

$$v^*A = \bar{\lambda}v^*.$$

Then, we can compute from both our equations:

$$\lambda v^*v = v^*\lambda v = v^*Av = \bar{\lambda}v^*v$$

and $v^*v = \sum |v_i|^2 > 0$, so

$$\lambda = \bar{\lambda}.$$

Thus, $\lambda \in \mathbb{R}$. □

Corollary

If A is a Hermitian matrix, then $\det A, \operatorname{tr} A \in \mathbb{R}$ too.

Proof. We can express

$$\det A = \text{ product of eigenvalues}$$
$$\operatorname{tr} A = \text{ sum of eigenvalues}$$

and since the eigenvalues are real, the determinant and trace are as well. □

Corollary

If $A \in R^{n \times n}$ is a symmetric matrix, then the eigenvalues of A are real.

Fact

To continue, we fix one of the settings in our following discussion:

- V is a $\mathbb{R}$-vector space equipped with a symmetric bilinear form $\langle\ ,\ \rangle$
- V is a $\mathbb{C}$-vector space equipped with a Hermitian form $\langle\ ,\ \rangle$

Definition (Orthogonality)

The notation $v \perp w$ means $\langle v, w \rangle = 0$. Orthogonality to a subspace of V, $v \perp W$, means $v \perp w$ for all $w \in W$. Between two subspaces, $W_1 \perp W_2$ means $w_1 \perp w_2$ for all $w_1 \in W_1$, $w_2 \in W_2$.

Example

Considering $V =$ spacetime $= (\mathbb{R}^4$ with $x_1x_2 + y_1y_2 + z_1z_2 - t_1t_2)$ and $v = (1, 0, 0, 1)$.

Then,

$$\langle v, v \rangle = 0$$

so $v \perp v$.

Definition (Orthogonal/Orthonormal Basis)

A basis $(v_1, \ldots, v_n)$ is an **orthogonal basis** if it is a basis where $v_i \perp v_j$ whenever $i \neq j$. Furthermore, this basis is an **orthonormal basis** if additionally, $\langle v_i, v_i \rangle = 1$ for all i (equivalently, $\langle v_i, v_j \rangle = \delta_{ij}$ for all i, j).

Definition (Orthogonal Space)

If $W \leq V$ is a subspace, then its **orthogonal space**

$$W^{\perp} := \{v \in V : v \perp W\}$$

is another subspace of V.

Definition (Nullspace)

The **nullspace/kernel** of a form $\langle\ ,\ \rangle$ is

$$V^{\perp} = \{v \in V : v \text{ is orthogonal to everything}\}$$

Example

Consider $V = \mathbb{R}^3$ with $\langle x, y\rangle := x_1y_1 + x_3y_3$. Then,

$$V^{\perp} = \{(0, a, 0) : a \in \mathbb{R}\}.$$

Example

Consider $V = \mathbb{R}^n$ where $\langle\ ,\ \rangle$ is identically zero. Then,

$$V^{\perp} = V.$$

Proposition 57

If $V = \mathbb{R}^n$ or $\mathbb{C}^n$ with $\langle x, y\rangle := x^*Ay$, then the kernel of $\ker\langle\ ,\ \rangle = \ker A$.

Proof. By definition,

$$y \in \ker\langle\ ,\ \rangle \iff \langle x, y\rangle = 0 \text{ for all } x.$$

Equivalently, for all x,

$$x^*Ay = 0$$

meaning that $Ay = 0$ (otherwise we could just choose x^* to return the ith value of Ay), and equivalently, $y \in \ker A$. □

Definition (Nondegeneracy)

The form $\langle\ ,\ \rangle$ is **nondegenerate** if $V^\perp = \{0\}$.

Example

Consider again $V =$ spacetime.

If $v = (a, b, c, d) \in V^\perp$, then pairing with each of the standard basis vectors should yield 0. According to the spacetime formula, this means that $a = b = c = d = -0$, and hence v is 0. Thus, $\langle\ ,\ \rangle$ is nondegenerate. Supposing that

$$W := \mathbb{R} \cdot (1, 0, 0, 1) \leq V$$

then $\langle\ ,\ \rangle$ on W is identically zero, which is *not* nondegenerate.

Definition (Nondegeneracy on a Subspace)

For a subspace $W \leq V$, $\langle\ ,\ \rangle$ is **nondegenerate on** W if the restriction of $\langle\ ,\ \rangle$ to a form on W is nondegerenate, or equivalently, the only $w \in W$ such that $w \perp W$ is $\vec{0}$. Fundamentally, $W \cap W^\perp = \{0\}$.

Proposition 58

Suppose $V = \mathbb{R}^n$ or $\mathbb{C}^n$ with $\langle x, y\rangle := x^* A y$. Then $\langle\ ,\ \rangle$ is nondegenerate if and only if A is invertible.

Proof. We can see that $\langle\ ,\ \rangle$ is nondegerenate if and only if $\ker\langle\ ,\ \rangle$ is $\{0\}$ if and only if $\ker A = \{0\}$ if and only if A is invertible. □

Example

We can consider $\langle\ ,\ \rangle$ on spacetime which is nondegerenate and

$$A = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} \text{ is invertible.}$$

Example

If $\langle\ ,\ \rangle$ is positive definite, then $\langle\ ,\ \rangle$ is nondegenerate (and nondegenerate on every subspace).

If $v \neq 0$, then $\langle v, v \rangle > 0$, so $v \notin V^\perp$.

§25 October 30, 2020

Theorem 59

Let $W \leq V$ be a subspace. The following conditions on $\langle\ ,\ \rangle$ are equivalent:

- $\langle\ ,\ \rangle$ is nondegenerate on W
- the only vector $w \in W$ such that $w \perp W$ is 0.
- $W \cap W^\perp = \{0\}$
- $V = W \oplus W^\perp$

Proof. Notice that $V = W \oplus W^\perp$ implies the third statement. Suppose then that $W \cap W^\perp = \{0\}$. Then, $W + W^\perp$ is a direct sum; we must determine whether this is all of V.

Choose a basis $(w_1, \ldots, w_k)$ of W, so $\dim W = k$. Define

$$\begin{aligned} \phi : V &\to \mathbb{C}^k \\ v &\mapsto (\langle w_1, v \rangle, \ldots, \langle w_k, v \rangle) \end{aligned}$$

Then $\ker \phi = W^\perp$, since each vector in the kernel is perpendicular to the basis vectors, and thus perpendicular to their span. By the rank-nullity theorem,

$$\dim V = \dim(\operatorname{im} \phi) + \dim(\ker \phi) \leq k + \dim W^\perp = \dim W + \dim W^\perp = \dim(W + W^\perp).$$

In particular,

$$\dim(W + W^\perp) \leq \dim V$$

and thus equality must hold everywhere, since there is $\dim V$ on both sides. Hence,

$$\dim(W \oplus W^\perp) = \dim V$$

and therefore,

$$W \oplus W^\perp = V.$$

□

Theorem 60

V has an orthogonal basis.

Proof. We induct on $\dim V$.

Case 1: There exists $v \in V$ such that $\langle v, v\rangle \neq 0$. In this case, letting $W := \text{Span}(v)$, then $\langle\ ,\ \rangle$ is nondegenerate on W, so

$$V = W \oplus W^{\perp}$$

and thus $v \in W$ and an orthogonal basis of $W^{\perp}$ (by the inductive hypothesis) form an orthogonal basis of V.

Case 2: $\langle v, v\rangle = 0$ for all $v \in V$. In this case, for all v, w,

$$\langle v + w, v + w\rangle = \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle$$

and hence,

$$\langle v, w\rangle + \langle w, v\rangle = 0.$$

In the symmetric case, $2\langle v, w\rangle = 0$, so $\langle v, w\rangle = 0$ for all v, w.

In the Hermitian case, $\langle v, w\rangle + \overline{\langle v, w\rangle} = 0$, so $\langle v, w\rangle$ is purely imaginary for all v, w. Then, $\langle v, iw\rangle = i\langle v, w\rangle$ are both imaginary, so $\langle v, w\rangle = 0$.

In both cases, $\langle\ ,\ \rangle$ is identically 0, so we can choose any basis. □

Corollary

V has an orthogonal basis $(v_1, \ldots, v_n)$ such that $\langle v_i, v_i\rangle$ is 1, -1, or 0 for each i.

Suppose (x_1, x_2) is an orthogonal basis (for a two-dimensional Hermitian space) with $\langle x_1, x_1\rangle = 3$ and $\langle x_2, x_2\rangle = 5$. Then, take $v_1 = \frac{1}{\sqrt{3}}x_1$ and $v_2 = \frac{1}{\sqrt{5}}x_2$, so $\langle v_1, v_1\rangle = 1$ and $\langle v_2, v_2\rangle = -1$.

Fact

Considering our above corollary

- a nondegenerate form has all 0s
- a positive definite form has all 1s (V has orthonormal basis)
- a positive semidefinite form has all 1s and 0s.

Theorem 61 (Sylvester's Law)

The number of 1s, -1s, and 0s are determined by V and $\langle\ ,\ \rangle$. The **signature** of $\langle\ ,\ \rangle$ is (#1s, #$-$1s).

Example

We can consider the signatures of the following examples:

- The dot product on $\mathbb{R}^n$ has signature $(n, 0)$
- The Lorentz form on $\mathbb{R}^4$ has signature $(3, 1)$

Corollary (Matrix Form)

For $A \in C^{n\times n}$ a Hermitian matrix, there exists $P \in GL_n(\mathbb{C})$ such that P^*AP is a diagonal matrix with entries 1, -1, or 0.

If A is positive definite, then

$$P^*AP = I$$

so

$$A = P^*P$$

for some $P \in GL_n(\mathbb{C})$ (the inverse of the previous P).

Definition (Euclidean Space)

A **Euclidean space** is a finite-dimensional $\mathbb{R}$-vector space equipped with a positive definite symmetric bilinear form (which has an orthonormal basis and hence is isomorphic to $\mathbb{R}^n$ equipped with the dot product).

Definition (Hermitian Space)

A **Hermitian space** is a finite-dimensional $\mathbb{C}$-vector space equipped with a positive definite Hermitian form (which is isomorphic to $\mathbb{C}^n$ with standard Hermitian form $\sum \bar{x}_i y_i$).

Example

Suppose V is nondegenerate, with orthogonal basis $v_1, \dots, v_n$. Given $x \in V$,

$$x = c_1 v_1 + \cdots + c_n v_n$$

for some scalars c_i. We can pair with v_1 to compute

$$\langle v_1, x \rangle = c_1 \langle v_1, v_1 \rangle + c_2 \langle v_1, v_2 \rangle + \cdots + c_n \langle v_1, v_n \rangle$$

so

$$c_1 = \frac{\langle v_1, x \rangle}{\langle v_1, v_1 \rangle}.$$

Definition (Orthogonal Projection)

Suppose $W \leq V$ and $\langle\ ,\ \rangle$ is nondegenerate on W. Then $V = W \oplus W^\perp$, so each v is $w + u$ for $w \in W$, $u \in W^\perp$. Then, **orthogonal projections** projects

$$\begin{aligned} \pi : V &\to W \\ v &\mapsto w \end{aligned}$$

The formula for π, assuming $w_1, \dots, w_k$ is an *orthogonal* basis of W, is

$$v = \underbrace{c_1 w_1 + \cdots + c_k w_k}_{\pi(v)} + u$$

where

$$c_i = \frac{\langle w_i, v \rangle}{\langle w_i, w_i \rangle}.$$

If $(w_1, \dots, w_k)$ is an ortho*normal* basis, then $c_i = \langle w_i, v \rangle$.

§26 November 2, 2020

Fact

If $\langle\ ,\ \rangle$ is positive definite, define $|v| := \sqrt{\langle v, v \rangle}$.

Algorithm (Gram-Schmidt Process)

Given a Euclidean space V with form $\langle\ ,\ \rangle$ and a basis $(v_1, \dots, v_n)$ of V, we can output an *orthonormal* basis $(w_1, \dots, w_n)$ of V.

We can represent $V_1 := \text{Span}(v_1)$, $V_2 := \text{Span}(v_1, v_2)$, ..., etc. First, we can define

$$w_1 := \frac{1}{|v_1|} v_1.$$

Then, (w_1) is an orthonormal basis for V_1. Then, we can define

$$t_2 := V_1^{\perp} \text{ component of } v_2 = v_2 - \text{proj}_{V_1} v_2 = v_2 - \langle w_1, v_2 \rangle w_1,$$

so $t_2 \perp w_1$. Then,

$$w_2 := \frac{1}{|t_2|} t_2,$$

so (w_1, w_2) is an orthonormal basis for V_2. Continuing similarly,

$$t_3 := V_2^{\perp} \text{ component of } v_3 = v_3 - (\langle w_1, v_3 \rangle w_1 + \langle w_2, v_3 \rangle w_2).$$

Then, $t_3 \perp w_1, w_2$ so

$$w_3 := \frac{1}{|t_3|} t_3,$$

and thus (w_1, w_2, w_3) is an orthonormal basis for V_3. We can repeat this process to generate an orthonormal basis of V_n, i.e. an orthonormal basis of V.

Fact

Consider, as our setting, the Hermitian spaces of V, $\langle\ ,\ \rangle_V$ and W, $\langle\ ,\ \rangle_W$.

Proposition 62

For each $\mathbb{C}$-linear transformation $T : V \to W$, there exists a *unique* linear transformation $T^* : W \to V$, the adjoint of T, such that

$$\langle Tv, w \rangle_W = \langle v, T^* w \rangle_V \text{ for all } v \in V, w \in W.$$

Proof. We choose an orthonormal basis for V and W to assume $V = \mathbb{C}^n$ with standard Hermitian form ($\langle x, y \rangle := x^* y$) and $W = \mathbb{C}^m$ with its standard Hermitian form. Then, T is given by some $A \in \mathbb{C}^{m \times n}$. We are looking for $B \in \mathbb{C}^{n \times m}$ such that

$$(Av)^* w = v^* B w \text{ for all } v \in V, w \in W.$$

Thus we can take $B = A^*$, and testing with $v = e_i, w = e_j$ shows that each entry of B must equal the corresponding entry of A^*. □

Fact

From now on, consider T as the linear operator $T : V \to V$, so $T^* : V \to V$.

Definition (Hermitian/Unitary/Normal Linear Operator)
Consider the following definitions:

- T is **Hermitian** $\iff T^* = T \iff \langle Tv, w\rangle = \langle v, Tw\rangle$ for all $v, w \in V$.
- T is **unitary** $\iff T^*T = I \iff \langle Tv, Tw\rangle = \langle v, w\rangle$ for all $v, w \in V$.
- T is **normal** $\iff T^*T = TT^* \iff \langle Tv, Tw\rangle = \langle T^*v, T^*w\rangle$ for all $v, w \in V$.

Both Hermitian and and unitary operators/matrices are special cases of normal ones.

Proposition 63
If $TW \leq W$, then $T^*W^\perp \leq W^\perp$.

Proof. Suppose that $u \in W^\perp$. If $w \in W$, then

$$\langle w, T^*u\rangle = \langle \underbrace{Tw}_{\in W}, \underbrace{u}_{\in W^\perp}\rangle = 0,$$

so $T^*u \in W^\perp$. □

Proposition 64
Suppose T is normal. If $Tv = \lambda v$, then $T^*v = \bar{\lambda}v$.

Proof. **If** $\lambda = 0$, then $Tv = 0$. Then,

$$\langle T^*v, T^*v\rangle = \langle Tv, Tv\rangle = \langle 0, 0\rangle = 0,$$

so $T^*v = 0$.

If λ is arbitrary, we let $S := T - \lambda I$. Then, $S^* = T^* - \bar{\lambda}I$, and

$$SS^* = S^*S,$$

since all terms $T, T^*, \lambda I, \bar{\lambda}I$ commute, so S is normal. Then,

$$Sv = Tv - \lambda v = 0$$

and by the $\lambda = 0$ case applied to S,

$$S^*v = 0 \implies T^*v = \bar{\lambda}v.$$

□

Theorem 65 (Spectral Theorem)

Given a Hermitian space V and a normal operator $T : V \to V$, then V has an orthonormal basis consisting of eigenvectors of T.

Matrix Version: Given a normal matrix $A \in C^{n\times n}$, then A is diagonalizable; more precisely, there exists unitary P (basechange matrix from standard basis to other orthonormal basis of $\mathbb{C}^n$) such that $P^{-1}AP$ is diagonal.

Proof. We induct on $\dim V$. Let w be an eigenvector of T. Without loss of generality, $|w| = 1$, and let $W = \text{Span}(w)$. Then,

$$V = W \oplus W^{\perp}.$$

Since $Tw = \lambda w$, both $TW \leq W$ and $T^*w = \bar{\lambda}w$ (Proposition 64), so $T^*W \leq W$, and hence $TW^{\perp} \leq W^{\perp}$ (Proposition 63). We have that (w) is an orthonormal basis for W, and induction yields an orthonormal basis of $W^{\perp}$ of eigenvectors of T. Put together, they will form an orthonormal basis of V consisting of eigenvectors of T. □

§27 November 4, 2020

Example

An example of a quadratic form in x, y, z is

$$10x^2 - 20y^2 + 30z^2 + 4xy + 6xz - 8yz = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 10 & 2 & 3 \\ 2 & -20 & -4 \\ 3 & -4 & 30 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Fact

More generally, a quadratic form in $x_1, \ldots, x_n$ is given by

$$\sum a_{ii}x_i^2 + \sum_{i<j} 2a_{ij}x_ix_j = \sum_{i,j} a_{ij}x_ix_j = x^tAx$$

where A is a real symmetric matrix (a_{ij}).

We can also perform an orthogonal coordinate change $x = Px'$, where $P \in O_n$. This changes A to P^tAP. By the spectral theorem for real symmetric matrices, we can make

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

with $\lambda_i \in \mathbb{R}$, getting a diagonal quadratic form

$$\lambda_1 x_1^2 + \cdots + \lambda_n x_n^2.$$

Fact

We can determine

$$\text{signature of the quadratic form} := (\#\text{ positive } \lambda_i, \#\text{ negative } \lambda_i)$$

Fact

Plane curves of degree 2 have form

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c = 0$$

in $\mathbb{R}^2$, or equivalently,

$$x^tAx + Bx + c = 0,$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \neq 0, B = \begin{pmatrix} b_1 & b_2 \end{pmatrix}.$$

Consider the possibilities

Possibilities	Examples
ellipse	$ax_1^2 + bx_2^2 - 1 = 0$
hyperbola	$ax_1^2 - bx_2^2 - 1 = 0$
parabola	$ax_1^2 - x_2 = 0$
two intersecting lines	$(ax_1 + bx_2)(cx_1 + dx_2) = 0,\ ad - bc \neq 0$
two parallel lines	$x_1^2 - a = 0,\ a > 0$
one line	$x_1^2 = 0$
one point	$x_1^2 + x_2^2 = 0$
empty	$x_1^2 + x_2^2 + 1 = 0$

where the first three cases are conic sections and the last five are degenerate cases.

Theorem 66

Every possibility is congruent (mappable by some isometry to) to one of the above examples.

Proof. Given $x^tAx + Bx + c = 0$, without loss of generality, $A = \begin{pmatrix} a & \\ & b \end{pmatrix}$ (diagonalized by Spectral Theorem), yielding the equation

$$ax_1^2 + bx_2^2 + b_1x_1 + b_2x_2 + c = 0.$$

Case 1: $a, b \neq 0$
We can complete the square by substituting $x_1 = x_1' - r$ for some $r \in \mathbb{R}$ to eliminate the b_1x_1 term, i.e., $r := \frac{b_1}{2a}$. We can do the same for x_2, yielding

$$ax_1^2 + bx_2^2 - c = 0$$

where c is a different constant. If $c \neq 0$, dividing the equation by c yields either an ellipse or a hyperbola. If $c = 0$, we get a point or two intersecting lines.

Case 2: $a \neq 0, b = 0$
We can proceed similarly with the same kind of argument. Remainder of proof left as exercise. □

Fact
We can also consider quadric surfaces (degree 2) in $\mathbb{R}^3$.

Similar analysis leads to

- $q(x_1, x_2, x_3) = 1$
- $x_3 = Q(x_1, x_2)$
- degenerate cases

Fact
The method used to describe conics can be applied to classify quadrics in any dimension. The general quadratic equation has the form $f = 0$, where

$$f(x_1, \ldots, x_n) = \sum_i a_{ii}x_i^2 + \sum_{i<j} 2a_{ij}x_ix_j + \sum_i b_ix_i + c.$$

Example
We can also consider skew-symmetric forms, where F is a field of characteristic not 2, and V is a finite dimensional F-vector space with bilinear form

$$\langle\ ,\ \rangle : V \times V \to F.$$

Proposition 67

For all x, y, $\langle x, y\rangle = -\langle y, x\rangle$ indicates that $\langle\ ,\ \rangle$ is **skew-symmetric**. If $\operatorname{char} F \neq 2$, the condition is equivalent to

$$\langle x, x\rangle = 0 \text{ for all } x, \text{ meaning } \langle\ ,\ \rangle \text{ is } \textbf{alternating}.$$

If $\operatorname{char} F = 2$, the alternating condition is more restrictive, and it is the better notion.

Proof. If $\langle x, y\rangle = -\langle y, x\rangle$ for all x, y, then

$$\langle x, x\rangle = -\langle x, x\rangle$$

so $\langle x, x\rangle = 0$ (supposing that $2 \neq 0$ in the field F). If $\langle x, x\rangle = 0$ for all x, then

$$\langle x + y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle$$

meaning that

$$0 = \langle x, y\rangle + \langle y, x\rangle$$

so

$$\langle x, y\rangle = -\langle y, x\rangle.$$

□

Example

An example of a skew-symmetric form on $\mathbb{R}^2$ is the "determinant" form:

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle := x_1 y_2 - x_2 y_1 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^t \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

which is alternating.

§28 November 6, 2020

Fact

Recall

- A real vector space V with a *symmetric* bilinear form $\langle\ ,\ \rangle$ has an *orthogonal* basis. (Same for a complex vector space with a *Hermitian* form.)
- If in addition $\langle\ ,\ \rangle$ is *positive definite*, then $\langle\ ,\ \rangle$ has an *orthonormal* basis.

Theorem 68

Let $\langle\ ,\ \rangle$ be a $\underbrace{\text{nondegenerate, alternating}}_{\text{symplectic}}$ form on V. Then, there exists a basis

$$(v_1, w_1, v_2, w_2, \ldots, v_n, w_n)$$

such that

$$\langle v_i, w_i \rangle = 1, \langle w_i, v_i \rangle = -1$$

and all other pairings of basis vectors give 0.

Matrix Version: $A \in GL_m(\mathbb{F})$ (invertible is equivalent to nondegenerate) such that

$$A^t = -A \text{ (with zeroes on the diagonal)}$$

Then, there exists some baschange matrix $P \in GL_m(\mathbb{F})$ such that

$$P^t A P = \begin{pmatrix} \sum & & \\ & \ddots & \\ & & \sum \end{pmatrix}$$

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

Proof. We proceed by induction on $\dim V$. If $\dim V = 0$, the empty basis satisfies all conditions. Otherwise, there exist v_1, w_1 with $\langle v_1, w_1 \rangle \neq 0$ (since $\langle\ ,\ \rangle$ is nondegenerate). Without loss of generality, scale w_1 so that $\langle v_1, w_1 \rangle = 1$. Since $\langle v_1, v_1 \rangle = 0$, $w_1 \notin \text{Span}(v_1)$; thus, let

$$W := \text{Span}(v_1, w_1) \text{ be a 2-dimensional space.}$$

In particular, $\langle\ ,\ \rangle$ is nondegenerate on W since its matrix is $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, which is invertible. Just as for symmetric forms,

$$V = W \oplus W^\perp$$

W has a basis (v_1, w_1) and by an inductive hypothesis, there exists a basis $(v_2, w_2, \ldots, v_n, w_n)$ of $W^\perp$. Concatenating these two bases yields a desired basis for their direct sum, V. $\square$

We can construct variants of the symplectic matrix by reording the basis.

Definition (Linear Group)

A **linear group** is a subgroup of $GL_n(\mathbb{R})$ or $GL_n(\mathbb{C})$ for some n.

Consider the following examples of linear groups:

Example (General Linear Group)

The **general linear group**

$$GL_n := GL_n(\mathbb{R})$$

Example (Orthogonal Group)

The **orthogonal group** (preserving dot product $x^t y$, length)

$$O_n := \{P \in GL_n(\mathbb{R}) : P^t P = I\}$$

Example (Lorentz Group)

The **Lorentz group** (preserving Lorentz form $x^t I_{3,1} y$)

$$O_{3,1} := \{P \in GL_n(\mathbb{R}) : P^t I_{3,1} P = I_{3,1}$$

where

$$I_{3,1} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}$$

Example (Unitary Group)

The **unitary group** (preserving standard Hermitian form $x^* y$)

$$U_n := \{P \in GL_n(\mathbb{C}) : P^* P = I\}$$

Example (Symplectic Group)

The **symplectic group** (preserve standard symplectic form $x^t S y$)

$$Sp_{2n} := \{P \in GL_{2n}(\mathbb{R}) : P^t S P = S\}$$

where

$$S = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

Example (Special Linear Group)
The **special linear group** (preserve oriented volume)

$$SL_n := \{P \in GL_n(\mathbb{R}) : \det P = 1\}$$

Example (Special Orthogonal Group)
The **special orthogonal group**

$$SO_n := \{P \in O_n : \det P = 1\}$$

Example (General Orthogonal Group)
The **general orthogonal group**

$$GO_n := \{P \in GL_n(\mathbb{R}) : \exists \lambda \in \mathbb{R}^\times \text{ such that } P^t P = \lambda I\}$$

Example (General Unitary Group)
The **general unitary group**

$$GU_n := \{P \in GL_n(\mathbb{R}) : \exists \lambda \in \mathbb{R}^\times \text{ such that } P^* P = \lambda I\}$$

Example (General Symplectic Group)
The **general symplectic group**

$$GSp_{2n} := \{P \in GL_{2n}(\mathbb{R}) : \exists \lambda \in \mathbb{R}^\times \text{ such that } P^t S P = \lambda S\}$$

Each of these groups has a geometry.

Example
$GL_n(\mathbb{R})$ is an open subset of $\mathbb{R}^{n\times n} \simeq \mathbb{R}^{n^2}$ since it is the inverse image under

$$\det : \mathbb{R}^{n\times n} \to \mathbb{R}$$

of the open subset $\mathbb{R}^\times \subset \mathbb{R}$. We say that $\dim GL_n(\mathbb{R}) = n^2$ (though it is not a vector space).

Example

The following are isomorphic:

$$\mathbb{R}/\mathbb{Z}, \mathbb{R}/2\pi\mathbb{Z}, U_1 := \{z \in \mathbb{C} : \bar{z}z = 1\}, SO_2$$

and is homeomorphic to the circle S^1.

Geometrically, they form a circle.

$$\begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto 2\pi x \end{cases}$$

is an isomorphism mapping $\mathbb{Z}$ to $2\pi\mathbb{Z}$, so $\mathbb{R}/\mathbb{Z} \simeq \mathbb{R}/2\pi\mathbb{Z}$.

$$\begin{cases} \mathbb{R} \twoheadrightarrow U_1 \\ \theta \mapsto e^{i\theta} \end{cases}$$

is a surjective homomorphism with kernel $2\pi\mathbb{Z}$ so it induces an isomorphism $\mathbb{R}/2\pi\mathbb{Z} \simeq U_1$. Similarly,

$$\begin{cases} \mathbb{R} \twoheadrightarrow SO_2 \\ \theta \mapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \end{cases}$$

is a surjective homomorphism with kernel $2\pi\mathbb{Z}$ so it induces an isomorphism $\mathbb{R}/2\pi\mathbb{Z} \simeq SO_2$.

§29 November 9, 2020

Definition (Homeomorphic)

Two topological spaces X, Y are **homeomorphic** ("have the same shape") if there exist continuous maps $f : X \to Y$ and $g : Y \to X$ such that $fg = 1_Y$ and $gf = 1_X$.

Fact

- In $\mathbb{R}^2$, the 1-sphere S^1: $x_0^2 + x_1^2 = 1$ and 2-ball B^2: $x_0^2 + x_1^2 \leq 1$.
- In $\mathbb{R}^3$, the 2-sphere S^2: $x_0^2 + x_1^2 + x_2^2 = 1$ and 3-ball B^3: $x_0^2 + x_1^2 + x_2^2 \leq 1$.
- In $\mathbb{R}^4$, the 3-sphere S^3: $x_0^2+x_1^2+x_2^2+x_3^2 = 1$ and 4-ball B^3: $x_0^2+x_1^2+x_2^2+x_3^2 \leq 1$.

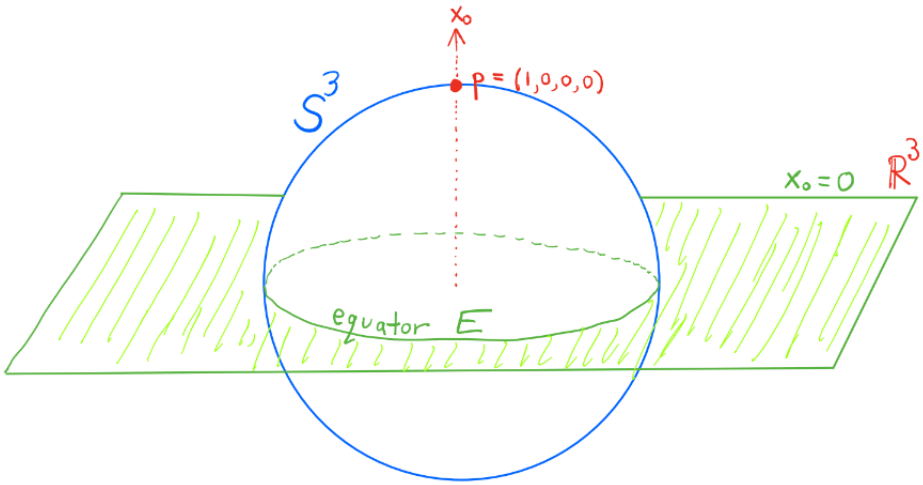
Theorem 69

The group

$$SU_2 = \{P \in GL_2(\mathbb{C}) : P^*P = I \text{ and } \det P = 1\}$$

is homemorphic to the 3-sphere S^3.

We can consider the 3-sphere



where the equator E is given by a 2-sphere.

Proposition 70

$$SU_2 = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} : a, b \in \mathbb{C}, \bar{a}a + \bar{b}b = 1 \right\}.$$

Proof. Let $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}^{2\times 2}$. Then

$$P \in SU_2 \iff P^*P = I \text{ and } \det P = 1,$$

if and only if $P^* = P^{-1}$ and $ad - bc = 1$, so

$$\begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \text{ and } ad - bc = 1.$$

Then, $d = \bar{a}$, $c = -\bar{b}$, and $\bar{a}a + \bar{b}b = 1$. □

Corollary

$$SU_2 = \left\{ x_0 \underbrace{\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}}_{\mathbb{1}} + x_1 \underbrace{\begin{pmatrix} i & \\ & -i \end{pmatrix}}_{\mathbb{i}} + x_2 \underbrace{\begin{pmatrix} & 1 \\ -1 & \end{pmatrix}}_{\mathbb{j}} + x_3 \underbrace{\begin{pmatrix} & i \\ i & \end{pmatrix}}_{\mathbb{k}} \right\}$$

where $x_0, \dots, x_3 \in \mathbb{R}$ and $x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1$ (from $a = x_0 + x_1 i$ and $b = x_2 + x_3 i$). This is the unit 3-sphere in the quaternion algebra

$$\mathbb{H} = \{x_0 \cdot \mathbb{1} + x_1 \cdot \mathbb{i} + x_2 \cdot \mathbb{j} + x_3 \cdot \mathbb{k} : x_0, x_1, x_2, x_3 \in \mathbb{R}\}$$

Fact

The Hamilton quaternions satisfy all the axioms of a field except that multiplication is not commutative.

Fact

The quaternion group of order 8

$$Q_8 := \{\pm\mathbb{1}, \pm\mathbb{i}, \pm\mathbb{j}, \pm\mathbb{k}\}$$

is a subgroup of SU_2.

Fact (About $P \in SU_2$)

Consider the facts about $P \in SU_2$:

- $\{\text{eigenvalues of } P\} = \{\lambda, \bar{\lambda}\}$ for some $\lambda \in \mathbb{C}$ with $|\lambda| = 1$
- characteristic polynomial of P is $t^2 - 2x_0 t + 1$ with $-1 \le x_0 \le 1$.
- P is conjugate in SU_2 to $\begin{pmatrix} \lambda & \\ & \bar{\lambda} \end{pmatrix}$
- $P, P' \in SU_2$ are conjugate in $SU_2 \iff \operatorname{tr} P = \operatorname{tr} P'$ (P and P' have the same x_0 value)

Proof. unitary implies eigenvalues have magnitude 1, and their product is the determinant of P, which is 1, so $\bar{\lambda} = \lambda^{-1}$. □

Proof. $\operatorname{tr} P = 2x_0$ and $\det P = 1$. □

Proof. P is unitary, so the spectral theorem says there exists $Q \in U_2$ such that

$$Q^{-1}PQ = \begin{pmatrix} \lambda & \\ & \bar{\lambda} \end{pmatrix}$$

Replace Q by αQ where $\alpha^2(\det Q) = \det(\alpha Q) = 1$. □

Proof. Conjugate matrices in $\mathbb{C}^{n\times n}$ have the same characteristic polynomial and hence the same trace. On the other hand, if $\operatorname{tr} P = \operatorname{tr} P'$, then the characteristic polynomials of P and P' are the same. Hence, the eigenvalues of P are the same as those of P', say $\{\lambda, \bar{\lambda}\}$. Then P and P' are both conjugate to $\begin{pmatrix} \lambda & \\ & \bar{\lambda} \end{pmatrix}$ and must be conjugate to each other. □

Definition (Latitude)

A **latitude** of c in SU_2 is the set of elements in SU_2 with $x_0 = c$ (a 2-sphere), where $x_0 = \frac{1}{2}\operatorname{tr} P$. Since two elements of SU_2 with the same trace are conjugate, each latitude is a conjugacy class in SU_2.

The center of $SU_2 = \{\pm\mathbb{1}\} = \{I, -I\}$, the north and south poles (single element in their latitudes, though the singleton sets are not necessarily considered latitudes).

Proposition 71

The **equator** is given by

$$E = \{P \in \mathbb{H} : P^2 = -\mathbb{1}\}$$

and is the latitude with $x_0 = 0$.

Proof. If $P \in E$, the characteristic polynomial of E is $t^2 + 1$, so $P^2 + \mathbb{1} = 0$ (by Cayley-Hamilton), so $P^2 = -\mathbb{1}$.

If $P^2 = -I$, then taking the determinant gives $(\det P)^2 = 1$, so $\det P = \pm 1$, meaning

$$x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1.$$

Thus, $P \in SU_2$ so P is conjugate to $\begin{pmatrix} \lambda & \\ & \bar{\lambda} \end{pmatrix}$, and since $P^2 = -\mathbb{1}$, P is

$$\pm \begin{pmatrix} i & \\ & -i \end{pmatrix}$$

so $\operatorname{tr} P = 0$, and hence $P \in E$. □

Example (Longitudes)

If $W \leq \mathbb{H}$ is a 2-dimensional subspace containing $\mathbb{1}$ and $-\mathbb{1}$ (i.e. a longitude), choose an orthonormal basis $\mathbb{1}, v$. Then $v \in E$, so $v^2 = -\mathbb{1}$, meaning

$$W = \{a \cdot \mathbb{1} + b \cdot v : a, b \in \mathbb{R}\} \text{ is a copy of } \mathbb{C}.$$

In particular, $W \cap SU_2$ is the unit circle in that copy of $\mathbb{C}$,

$$\{(\cos\theta)\mathbb{1} + (\sin\theta)v : \theta \in [0, 2\pi)\},$$

which is a subgroup of SU_2.

Fact

Also,

- The longitudes are conjugate subgroups, because the v's are conjugate
- Their union is SU_2, because every $P \in SU_2$ belongs to some W.

§30 November 13, 2020

Fact

Let $\mathbb{V} := \ker(\text{tr} : \mathbb{H} \to \mathbb{R}) = \text{Span}(\mathbb{i}, \mathbb{j}, \mathbb{k}) \simeq \mathbb{R}^3$ (essentially copy of $\mathbb{R}^3$), which we can equip with a dot product to get a symmetric bilinear form on $\mathbb{V}$, with $\mathbb{i}, \mathbb{j}, \mathbb{k}$ as an orthonormal basis.

If $u, v \in \mathbb{V}$, then

$$\begin{aligned} uv &= (u_1\mathbb{i} + u_2\mathbb{j} + u_3\mathbb{k})(v_1\mathbb{i} + v_2\mathbb{j} + v_3\mathbb{k}) \\ &= (-u \cdot v)\mathbb{1} + (u \times v) \end{aligned}$$

so

$$u \cdot v = -\frac{1}{2}\,\text{tr}(uv).$$

Definition (Conjugation by P)

For $P \in SU_2$, we can define "conjugation by P" which maps

$$\begin{aligned} \mathbb{H} &\to \mathbb{H} \\ x &\mapsto PxP^{-1} \end{aligned}$$

This preserves trace, so it restricts to a map

$$\begin{aligned} \gamma_P : \mathbb{V} &\to \mathbb{V} \\ x &\mapsto PxP^{-1} \end{aligned}$$

Theorem 72

We have the following set of properties:

(1) $\gamma_P \in SO_3$. More specifically, if $P = (\cos\theta)\mathbb{1} + (\sin\theta)v$ for some $v \in E$, then γ_P is rotation by 2θ about the pole v.

(2) There is a surjective homomorphism

$$\begin{aligned} SU_2 &\to SO_3 \\ P &\mapsto \gamma_P \end{aligned}$$

with kernel $\{\pm I\}$.

Proof. Since v is conjugate to $\mathbb{i}$ (both are in the conjugacy class E), by symmetry we may assume $v = \mathbb{i}$, so

$$P = (\cos\theta)\mathbb{1} + (\sin\theta)\mathbb{i}.$$

Then,

$$\gamma_P(\mathbb{i}) = P\mathbb{i}P^{-1} = \mathbb{i}.$$

We can also consider

$$\gamma_P(\mathbb{j}) = P\mathbb{j}P^{-1} = (\cos 2\theta)\mathbb{j} + (\sin 2\theta)\mathbb{k}$$

and

$$\gamma_P(\mathbb{k}) = P\mathbb{k}P^{-1} = (-\sin 2\theta)\mathbb{j} + (\cos 2\theta)\mathbb{k}.$$

While $\mathbb{i}$ is preserved, a rotation by 2θ occurs in the $(\mathbb{j}, \mathbb{k})$-plane. Furthermore, every element of SO_3 is a rotation of the type in (1), so $SU_2 \to SO_3$ is surjective. Also,

$$\gamma_P = I \iff 2\theta \in 2\pi\mathbb{Z} \iff P = \pm I.$$

□

Corollary

We can think about

$$SO_3 \approx \{\text{cosets of } \{\pm I\} \text{ in } SU_2\} \approx S^3/\sim,$$

where $v \sim w \iff v = \pm w$.

We hop back to Chapter 5 material to discuss differential equations and the matrix exponential.

Fact

Fix $a \in \mathbb{C}$; the function $x(t) = e^{at}$ for $t \in \mathbb{R}$ is the unique solution to $(\star)$, the differential equation

$$\frac{dx}{dt} = ax$$

with initial condition $x(0) = 1$.

We have two ways to construct this function:

(1) Invoke the general "existence and uniqueness theorem for linear ordinary differential equations"

(2) Prove that

$$1 + at + \frac{(at)^2}{2!} + \frac{(at)^3}{3!} + \cdots$$

converges to a differentiable function satisfying $(\star)$

Example

Define $a := i$. Then

$$e^{it} = \cos t + i \sin t$$

because the right-hand side satisfies $(\star)$.

We can consider approach (2) for $A \in \mathbb{C}^{n \times n}$.

Definition (Exponential of Matrix)

Define

$$e^A := I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

so we have the matrix valued function

$$e^{tA} = I + tA + \frac{(tA)^2}{2!} + \frac{(tA)^3}{3!} + \cdots$$

We need to check

- this converges for every $t \in \mathbb{R}$
- is differentiable
- $\frac{d}{dt}e^{tA} = Ae^{tA}$
- $e^0 = I$

In general, suppose $u_1(t), u_2(t), \ldots$ are functions from $I \to \mathbb{R}$ (where I is some interval in $\mathbb{R}$). We would like to show

1. The sum $\sum_k u_k(t)$ converges uniformly if there exist real numbers M_k such that $|u_k(t)| \leq M_k$ for all $t \in I$ and $\sum M_k < \infty$ (Weierstrass M-test)
2. If each u_k is continuous and $\sum u_k$ converges uniformly, then
$$u(t) := \sum_k u_k(t) \text{ is continuous}$$
3. If each u_k is differentiable and $\sum u_k(t)$ converges and $\sum u_k'(t)$ converges uniformly, then u is differentiable and
$$u'(t) = \sum_k u_k'(t)$$

We can define $\|A\| := \max_{i,j} |a_{ij}|$. By induction on k, we have $\left\|A^k\right\| \leq n^{k-1} \|A\|^k$. Let $I = (-r, r)$ and $V = \mathbb{C}^{n\times n}$. To define e^{tA}, we want to sum

$$u_k(t) := \frac{(tA)^k}{k!}, \text{ which has } u_k'(t) = \frac{t^{k-1}A^k}{(k-1)!}.$$

These are bounded on $(-r, r)$ by

$$M_k := \frac{r^k n^{k-1} \|A\|^k}{k!} \text{ and } N_k := \frac{r^{k-1} n^{k-1} \|A\|^k}{(k-1)!}.$$

Then $\sum u_k(t)$ converges uniformly on $(-r,r)$ since $\sum M_k < \infty$ and $\sum u_k'(t)$ converges uniformly on $(-r,r)$ since $\sum N_k < \infty$. Then,

$$e^{tA} := I + tA + \frac{(tA)^2}{2!} + \frac{(tA)^3}{3!} + \cdots$$

converges uniformly on $(-r,r)$ and we can compute

$$\frac{d}{dt}e^{tA} = Ae^{tA},$$

at least for $t \in (-r,r)$.

§31 November 16, 2020

Fact

Consider some properties of e^{tA}:

- $\frac{d}{dt}e^{tA} = Ae^{tA}$ for all $t \in \mathbb{R}$ (true on each bounded interval)
- $e^0 = I$
- If $AB = BA$, then $e^Ae^B = e^{A+B}$
- e^A is inveritble, with inverse e^{-A}
- If $B = PAP^{-1}$, then $e^B = Pe^AP^{-1}$.
- If $A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$, then $e^A = \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix}$

Most of these are relatively straightforward. We can prove that $AB = BA$ implies $e^Ae^B = e^{A+B}$. In particular, if $AB = BA$, then the binomial theorm applies:

$$(A+B)^m = \sum_{k+l=m} \frac{m!}{k!}l!A^kB^l.$$

Then, by definition,

$$\begin{aligned} e^{A+B} = \sum_{m\geq 0}\frac{(A+B)^m}{m!} &= \sum_{m\geq 0}\sum_{k+l=m}\frac{1}{m!}\frac{m!}{k!l!}A^kB^l \\ &= \sum_{k\geq 0}\sum_{l\geq 0}\frac{A^k}{k!}\frac{B^l}{l!} \\ &= e^Ae^B. \end{aligned}$$

In this case, it is okay to rearrange terms because everything converges absolutely.

Corollary

For each $A \in \mathbb{C}^{n\times n}$,

$$\begin{aligned} \mathbb{R} &\to GL_n(\mathbb{C}) \\ t &\mapsto e^{tA} \end{aligned}$$

is a homomorphism.

In particular,

$$e^{(t+u)A} = e^{tA+uA} = e^{tA}e^{uA}$$

since tA and uA commute.

Definition (One-Parameter Group)

A **one-parameter group** in $GL_n(\mathbb{C})$ is a differentiable homomorphism

$$\phi : \mathbb{R} \to GL_n(\mathbb{C}).$$

- Each $A \in \mathbb{C}^{n\times n}$ gives a one-parameter group $\phi(t) := e^{tA}$.
- Every one-parameter group ϕ is e^{tA} for some $A \in \mathbb{C}^{n\times n}$, namely $A = \phi'(0)$.

Proposition

There is a bijection

$$\mathbb{C}^{n\times n} \leftrightarrow \{\text{one-parameter groups in } GL_n(\mathbb{C})\}$$

where

$$A \mapsto (t \mapsto e^{tA}).$$

In particular the one-parameter group ϕ corresponds to the complex matrix $A = \phi'(0)$.

Proof. We already showed that for each A, $t \mapsto e^{tA}$ is a one-parameter group. Suppose that $\phi : \mathbb{R} \to GL_n(\mathbb{C})$ is *any* one-parameter group and let $A = \phi'(0)$. We need to show that $\phi(t) = e^{tA}$ for all $t \in \mathbb{R}$. It is enough to show that $\phi(t)$ satisfies the same differential equation and initial condition as e^{tA}.

$$\phi(s+t) = \phi(s)\phi(t)$$

so taking the derivative with respect to s yields

$$\phi'(s+t) = \phi'(s)\phi(t)$$

and evaluating at $s = 0$ yields

$$\phi'(t) = A\phi(t),$$

which is the same differential equation that e^{tA} satisfies. Furthermore,

$$\phi(0) = I$$

since ϕ is a homomorphism. □

Example

Consider $A = \mathbb{i} = \begin{pmatrix} i & \\ & -i \end{pmatrix}$. Then

$$e^{tA} = \begin{pmatrix} e^{it} & \\ & e^{-it} \end{pmatrix} \in SU_2$$

is a one-parameter group that corresponds to the longitude through $\mathbb{i}$.

Example

For which $A \in \mathbb{C}^{n\times n}$ does the one-parameter group $t \mapsto e^{tA}$ land in $GL_n(\mathbb{R})$?

We claim that $A \in \mathbb{R}^{n\times n} \iff e^{tA} \in GL_n(\mathbb{R})$ for all $t \in \mathbb{R}$. If $A \in \mathbb{R}^{n\times n}$, then by definition, e^{tA} takes real values. On the other hand, if $e^{tA} \in GL_n(\mathbb{R})$ for all t, taking the derivative at $t = 0$ gives $A \in \mathbb{R}^{n\times n}$.

Example

For which $A \in GL_n(\mathbb{C})$ is $e^{tA} \in U_n$ for all $t \in \mathbb{R}$?

We claim that $\underbrace{A^* = -A}_{\text{skew-Hermitian}} \iff e^{tA} \in U_n$ for all $t \in \mathbb{R}$. In general,

$$(e^A)^* = e^{A^*}$$

by expanding both with the series definitions. If $A^* = -A$, then

$$(e^A)^* = e^{A^*} = e^{-A} = (e^A)^{-1}$$

so $e^A \in U_n$. Conversely, if $e^{tA} \in U_n$ for all $t \in \mathbb{R}$, then

$$(e^{tA})^* = (e^{tA})^{-1}$$

and thus

$$e^{tA^*} = e^{-tA}.$$

Taking the derivative with respect to t and setting $t = 0$ yields

$$A^* = -A.$$

§32 November 18, 2020

Example

For which A is $e^{tA} \in O_n$ for all $t \in \mathbb{R}$?

Since $O_n = GL_n(\mathbb{R}) \cap U_n$,

$$A \in \mathbb{R}^{n \times n} \text{ and } A^* = -A,$$

from the two examples from last lecture, meaning A is a real skew-symmetric matrix.

Lemma

For any $A \in \mathbb{C}^{n \times n}$, $\det e^A = e^{\operatorname{tr} A}$.

Proof. Conjugating A by P conjugates e^A, so it doesn't change $\det e^A$ and it also doesn't change the trace, $\operatorname{tr} A$. Without loss of generality, A is in Jordan canonical form (consider it as simply being upper diagonal) with diagonal entries $\lambda_1, \dots, \lambda_n$. Then, A^k has diagonal entries $\lambda_1^k, \dots, \lambda_n^k$ and is still upper triangular. Furthermore, e^A has diagonal entries $e^{\lambda_1}, \dots, e^{\lambda_n}$ and is upper triangular. Then,

$$\det e^A = e^{\lambda_1} \cdots e^{\lambda_n} = e^{\lambda_1 + \cdots + \lambda_n} = e^{\operatorname{tr} A}.$$

□

Example

For which $A \in \mathbb{R}^{n \times n}$ is $e^{tA} \in SL_n(\mathbb{R})$ for all $t \in \mathbb{R}$?

We claim that this is true for the A such that $\operatorname{tr} A = 0$. In particular, for $A \in \mathbb{R}^{n \times n}$, the following are equivalent:

- $\operatorname{tr} A = 0$
- $\operatorname{tr} tA = 0$ for all $t \in \mathbb{R}$
- $e^{\operatorname{tr} tA} = 1$ for all t
- $\det e^{tA} = 1$ for all t (by the above lemma)

Consequently, $e^{tA} \in SL_n(\mathbb{R})$ for all t.

Definition (Manifolds)

A ***d*-dimensional manifold** is a topological/metric space M (think of a subset of $\mathbb{R}^n$) such that every point $m \in M$ has an open neighborhood in M that is homeomorphic to an open subset of $\mathbb{R}^d$.

Example

A circle is a 1-dimensional manifold.

Every neighborhood of a point on the circle is homeomorphic to an interval in $\mathbb{R}^1$.

Example

A sphere is a 2-dimensional manifold.

Example

A 3-sphere is a 3-dimensional manifold]

Example

If U is an open set in $\mathbb{R}^d$ and $f : U \to \mathbb{R}^e$ is a continuous function, then

$$graph(f) \subset U \times \mathbb{R}^e \subset \mathbb{R}^{d+e}$$

is homeomorphic to U, where $graph(f)$ contains all points $(u, f(u))$ for $u \in U$.

We have that $U \to graph(f)$, $u \mapsto (u, f(u))$, and point (u, v) maps back to u, so $graph(f)$ is a d-dimensional manifold. This explains why the circle is a manifold – it's locally the graph of a function.

Fact

As a warmup question, when can you solve for y as a function of x in the equation

$$\underbrace{ax + by + c}_{f(x,y)} = 0.$$

That is, when does this equation describe the graph of a function?

We can solve for y in this case whenever $b \neq 0$. In other words, we need

$$\frac{\partial f}{\partial y} \neq 0.$$

Theorem 73 (Implicit Function Theorem)

Given

(1) a system of equations
$$f(x, y) = 0$$
where f is from a tuple $(f_1, \ldots, f_r)$, x is from $(x_1, \ldots, x_m)$, and y is from $(y_1, \ldots, y_r)$ where $f : U \to \mathbb{R}^r$ is a C^1 function (has continuous partial derivatives) and U is an open subset of $\mathbb{R}^{m+n}$

(2) a point $u \in \{f = 0\}$ (set of solutions to $f = 0$)

For
$$\left(\frac{\partial f}{\partial y}\right) = \left(\left(\frac{\partial f_i}{\partial y_j}\right)(u)\right) \in \mathbb{R}^{r \times r},$$
if $\left(\frac{\partial f}{\partial y}\right)(u)$ is invertible, then $\{f = 0\}$ is a graph near u, i.e., there is an open neighborhood U' of u in U such that the intersection of the solution set and the open neighborhood is $\{f = 0\} \cap U' = graph(g)$ for some C^1 function $g : V \to \mathbb{R}^r$ where V is an open subset of $\mathbb{R}^m$.

Example

We can prove that
$$S^2 : x^2 + y^2 + z^2 - 1 = 0 \text{ in } \mathbb{R}^3$$
is a manifold.

We need to check that for each point $u \in S^2$, an open neighborhood is homeomorphic to an open subset of $\mathbb{R}^2$. At u¡ not all of x, y, z can be 0, so without loss of generality, $z \neq 0$ at u (it is symmetric otherwise). We try to solve for z:
$$\frac{\partial f}{\partial z} = 2z \text{ is nonzero at } u,$$
so the implicit function theorem says that some neighborhood of u in S^2 is the graph of a function $g(x, y)$. Intuitively, this makes sense since we could write
$$g(x, y) = \sqrt{1 - x^2 - y^2} \text{ or } -\sqrt{1 - x^2 - y^2}.$$
In conclusion, S^2 is a manifold.

§33 November 20, 2020

Theorem 74

Let G be a closed subgroup (closed subset and subgroup) of $GL_n(\mathbb{R})$ (or $GL_n(\mathbb{C})$). Then G is a manifold; in fact, G is $\{f = 0\}$ for some function f satisfying the conditions of the Implicit Function Theorem at each point (choice of variables depends on the point).

Definition (Differentiable Path)

A **differentiable path** in $G \leq GL_n(\mathbb{R})$ as above is a differentiable function

$$\phi : \mathcal{I} \to G$$

where $\mathcal{I}$ is an open interval in $\mathbb{R}$. If $0 \in \mathcal{I}$, we get the **velocity vector** $\phi'(0) \in \mathbb{R}^{n \times n}$.

We can look at three ways to describe the Lie algebra $\mathfrak{g} = \text{Lie } G$:

Definition (Lie Algebra)

It's the *tangent space* to G at I

$$:= \{\phi'(0) : \phi \text{ is a differentiable path in } G \text{ with } \phi(0) = I\}$$

Definition (Lie Algebra)

It's

$$\{A \in \mathbb{R}^{n \times n} : \text{the one-parameter group } e^{tA} \text{ is contained in } G\}$$

Suppose A is in this set. Then, the one-parameter group $\phi(t) := e^{tA}$ is contained in G, so $\phi'(0) = A$ is in the set given by the first definition of a Lie Algebra.

Example (Algebraic Method of Derivation)

Working in $\mathbb{R}[\varepsilon] := \mathbb{R} + \mathbb{R}\varepsilon$ where $\varepsilon^2 = 0$, we might consider $f(x) = x^2 + 5x$. Then,

$$f(x + \varepsilon) = (x + \varepsilon)^2 + 5(x + \varepsilon) = (x^2 + 5x) + (2x + 5)\varepsilon.$$

Definition (Lie Algebra)

If G is defined by a system of polynomial equations $f = 0$ satisfying the Implicit Function Theorem, the Lie Algebra is

$$\{A \in \mathbb{R}^{n\times n} : I + \varepsilon A \text{ satisfies the system of equations}\}$$

Example ($\mathfrak{o}_n = \operatorname{Lie} O_n$)

Consider

$$O_n = \{P \in GL_n(\mathbb{R}) : P^t P = I\}.$$

Then

$$\operatorname{Lie} O_n = \{A \in \mathbb{R}^{n\times n} : (I + \varepsilon A)^t (I + \varepsilon A) = I\}$$

Consider

$$(I + \varepsilon A^t)(I + \varepsilon A) = I,$$

so

$$I + \varepsilon A^t + \varepsilon A = I,$$

meaning

$$\varepsilon(A^t + A) = 0,$$

and hence

$$A^t = -A,$$

meaning $\operatorname{Lie} O_n = \{A \in \mathbb{R}^{n\times n} : A^t = -A\}$.

Example ($\mathfrak{sl}_2 = \operatorname{Lie} SL_2$)

Consider

$$SL_2 = \{\det = 1\}.$$

Then,

$$\operatorname{Lie} SL_2 = \{A \in \mathbb{R}^{2\times 2} : \operatorname{tr} A = 0\}$$

Consider the following equivalent statements:

- $A \in \operatorname{Lie} SL_2$
- $\det(I + \varepsilon A) = 1$
- $\det \begin{pmatrix} 1 + \varepsilon a & \varepsilon b \\ \varepsilon c & 1 + \varepsilon d \end{pmatrix} = 1$
- $(1 + \varepsilon a)(1 + \varepsilon d) - (\varepsilon b)(\varepsilon c) = 1$

- $1 + \varepsilon a + \varepsilon d = 1$
- $\varepsilon(a + d) = 0$
- $\operatorname{tr} A = 0$

Fact (Properties of $\operatorname{Lie} G$)

Consider the following properties of $\operatorname{Lie} G$:

- $\operatorname{Lie} G$ is a vector space of the same dimension as G (as a manifold)
- If $A, B \in \operatorname{Lie} G$, then $AB - BA \in \operatorname{Lie} G$

For a sketch of the proof of the second property, use $\mathbb{R}[\varepsilon, \varepsilon']$ with $\varepsilon^2 = (\varepsilon')^2 = 0$, consisting of

$$\{a + b\varepsilon + c\varepsilon' + d\varepsilon\varepsilon' : a, b, c, d \in \mathbb{R}\}.$$

If $A, B \in \operatorname{Lie} G$, then $I + \varepsilon A, I + \varepsilon' B$ satisfy the equations of G and then so does their commutator

$$(I + \varepsilon A)(I + \varepsilon' B)(I + \varepsilon A)^{-1}(I + \varepsilon' B)^{-1} = I + (AB - BA)\varepsilon\varepsilon'$$

using that $(I + \varepsilon A)^{-1} = I - \varepsilon A$. In particular, $\varepsilon\varepsilon'$ is just another infinitesimal with square 0, so $AB - BA \in \operatorname{Lie} G$.

Definition (Bracket)

We define the **bracket**

$$[A, B] := AB - BA$$

which measures noncommutativity in some infinitesimal sense.

§34 November 30, 2020

Definition (Lie Algebra)

A **Lie Algebra** is a vector space V equipped with a binary operation

$$[\ ,\ ] : V \times V \to V$$

satisfying

- bilinearity
- $[A, A] = 0$ for all $A \in V$ (so $[B, A] = -[A, B]$)
- **Jacobi identity**:

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \text{ for all } A, B, C$$

A key idea is that groups like G are manifolds, but we can reduce to Lie Algebras, which are simply vector spaces, and simpler to deal with.

Example (Translation Map)

Given $g \in G$ (a linear group, i.e., closed subgroup of GL_n), the "translation map" $G \to G$, $x \mapsto gx$ is usually not a homomorphism (unless $g = 1$), but it is continuous and its inverse $x \mapsto g^{-1}x$ is continuous too, so it's a homeomorphism, and it sends 1 to g.

From the point of view of topology, "every point of G looks the same," a property known as homogeneity.

Example

By definition, G being a d-dimensional manifold is equivalent to saying every $g \in G$ has an open neighborhood homeomorphic to an open subset of $\mathbb{R}^d$. By homogeneity, it's enough to check this for $g = 1$.

Fact

Any closed subgroup $G \leq GL_n(\mathbb{R})$ is a manifold of the same dimension as its Lie algebra $\mathfrak{g}$

For an idea of the proof, we can consider the map

$$\begin{aligned} exp : \mathbb{R}^{n \times n} &\to GL_n(\mathbb{R}) \\ A &\mapsto e^A \end{aligned}$$

which homeomorphically maps a small neighborhood of 0 to a small neighborhood of I. It restricts to a differentiable map

$$\mathfrak{g} \to G$$
$$A \mapsto e^A$$

Proposition

If H is an open subgroup of a linear group G, then H is closed in G.

Proof. If we consider the cosets of H, which form all of G, then the homeomorphism $x \mapsto gx$ maps H to gH, but then gH ought to be open as well, since H is. Similarly, all cosets of H ought to be open. Taking the union of all cosets of H, except H itself, the union $G \backslash H$ ought to be open. Thus, H is closed. □

Example

An example of the above proposition is if $G = \mathbb{R}^\times$ and $H = \mathbb{R}_{>0}$.

Definition

G is **path-connected** if for every $x, y \in G$, there exists a continuous map

$$\phi : [0, 1] \to G$$

with $\phi(0) = x$ and $\phi(1) = y$.

Fact

If G is path-connected, then G is not a disjoint union of two nonempty open subsets.

Proposition

Suppose G is a path-connected linear group. Any nonempty open subset U of G generates the whole group G.

Proof. Let H be the subgroup generated by U. We want to show that $H = G$. Choose some $g \in U$. Then $V := g^{-1}U$ is an open neighborhood of 1 in G and is contained in H. For each $h \in H$, hV is an open neighborhood of h in G and is contained in H. Thus, H is open. G is a disjoint union of (open) cosets of H, but G is path-connected, so there's only one coset, meaning $H = G$. □

Fact (Simple Group)

Recall that G is *simple* if and only if $G \neq \{1\}$ and the only normal subgroups are $\{1\}$ and G.

Theorem 75

SU_2 is "almost simple": its only normal subgroups are $\{I\}$, $\{\pm I\}$, SU_2.

Proof. Let $N \triangleleft SU_2$ and suppose N is not $\{I\}$ or $\{\pm I\}$. Then, N contains some $g \neq \pm I$. Since N is normal, N contains the conjugacy class of g, a latitude Λ. Then N contains $g^{-1}\Lambda$, which passes through I. Then, N must contain the latitudes of all points in $g^{-1}\Lambda$, so N contains an open neighborhood of I. By the above proposition, $N = SU_2$. □

§35 December 2, 2020

Corollary

SO_3 is simple.

Proof. We have the spin homomorphism $\gamma : SU_2 \twoheadrightarrow SO_3$ with kernel $\{\pm I\}$. The Correspondence Theorem says

$$\{\text{normal subgroups of } SU_2 \text{ containing } \{\pm I\}\} \leftrightarrow \{\text{normal subgroups of } SO_3\}$$

Thus $SU_2 \leftrightarrow SO_3$ and $\{\pm I\} \leftrightarrow \{I\}$. □

Theorem 76

Let F be a field with $|F| \geq 4$. Then, the only normal subgroups of $SL_2(F)$ are $\{I\}$, $\{\pm I\}$, and $SL_2(F)$.

Note that in the case the characteristic of F is 2 (e.g., $F = \mathbb{F}_{2^n}$), then $-1 = 1$, so $-I = I$, so $SL_2(F)$ *is* actually simple.

Proof. We prove the theorem for $|F| > 5$. First, we utilize a lemma:

Lemma ($\sqrt{}$ Lemma)

For $a \in \mathbb{F}$, the equation $x^2 = a$ has ≤ 2 solutions.

If $x^2 = y^2 = a$, then $(x+y)(x-y) = 0$, so $y = \pm x$. In particular, if $|F| > 5$, then there exists $r \in F$ such that $r^2 \notin \{0, 1, -1\}$. In particular, $r^2 = 0$ has one solution, $r^2 = 1$ has ≤ 2 solutions, and $r^2 = -1$ has ≤ 2 solutions. Thus, we can choose an $r \in F$ that is not any of these.

For the proof of the theorem, suppose $N \lhd SL_2(F)$ and $N \not\subset \{\pm I\}$. We show that $N = SL_2(F)$:

Step 1: N contains some matrix B with distinct eigenvalues s, s^{-1}. We prove this by choosing $A \in N \backslash \{\pm I\}$. Then A is *not* a scalar multiple of I. Thus, there exists v_1 such that v_1 is *not* an eigenvector of A, so letting $v_2 = Av_1$, (v_1, v_2) is a basis of F^2 (linearly independent vectors). Choose $r \in F$ such that $r^2 \notin \{0, 1, -1\}$. Let $P \in GL_2(F)$ be such that

$$Pv_1 = rv_1 \text{ and } Pv_2 = r^{-1}v_2.$$

Then, $\det P = r \cdot r^{-1} = 1$, so $P \in SL_2(F)$. Since N is closed under conjugation,

$$PA^{-1}P^{-1} \in N \text{ and } B := APA^{-1}P^{-1} \in N.$$

Applying B to v_2, we get

$$v_2 \xrightarrow{P^{-1}} rv_2 \xrightarrow{A^{-1}} rv_1 \xrightarrow{P} r^2 v_1 \xrightarrow{A} r^2 v_2.$$

Thus, $Bv_2 = r^2 v_2$. Hence, the eigenvalues of B are r^2 and r^{-2}, since $\det B = 1$. Furthermore $r^2 \neq r^{-2}$, since $r^2 \neq \pm 1$.

Step 2: Let $C := \{\text{matrices in } SL_2(F) \text{ with eigenvalues } s, s^{-1}\}$. Then C is a single conjugacy class of $SL_2(F)$ (the conjugacy class of B). Let $S = \begin{pmatrix} s & \\ & s^{-1} \end{pmatrix}$. Then the conjugacy class of S is contained in C. Suppose that $Q \in C$. Then Q is conjugate in $GL_2(F)$ to S, i.e.,

$$Q = LSL^{-1}$$

for some $L \in GL_2(F)$. Then, we can write

$$Q = \underbrace{L \begin{pmatrix} a & \\ & 1 \end{pmatrix}}_{M} S \underbrace{\begin{pmatrix} a & \\ & 1 \end{pmatrix}^{-1} L^{-1}}_{M^{-1}}$$

for any $a \in F^\times$, since diagonal matrices commute. Thus, we can write

$$Q = MSM^{-1},$$

and we can choose $a = (\det L)^{-1}$, so that $\det M = 1$, so $M \in SL_2(F)$. Hence, Q is also conjugate in $SL_2(F)$.

Step 3: $\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \in N$ for all $x \in F$. In particular, we can write

$$\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} s^{-1} & \\ & s \end{pmatrix}}_{\in C \subset N} \underbrace{\begin{pmatrix} s & sx \\ & s^{-1} \end{pmatrix}}_{\in C \subset N} \in N$$

Step 4: We can similarly demonstrate that $\begin{pmatrix} 1 & \\ y & 1 \end{pmatrix} \in N$ for all $y \in F$.

Step 5: $N = SL_2(F)$, since the elementary matrices in Steps 3 and 4 generate $SL_2(F)$. □

Fact (Center of Special Linear Group)
We can determine $Z :=$ center of $SL_2(F)$.

In particular,

$$\{\pm I\} \subseteq Z \lhd SL_2(F)$$

and in particular, $Z \neq SL_2(F)$, since it is not abelian, and thus $Z = \{\pm I\}$.

Definition (Projective Special Linear Group)
For any field F, we can define a new group

$$PSL_2(F) := SL_2(F)/Z.$$

Notice that if $\operatorname{char} F = 2$, then $Z = \{I\}$, so $PSL_2(F) = SL_2(F)$.

Corollary
If $|F| \geq 4$, then $PSL_2(F)$ is simple.

Proof. We can prove this similarly to the first corollary of the day using the Correspondence Theorem. □

Example
We arrive at some "accidental" isomorphisms.

We can consider the isomorphisms

q	$PSL_2(\mathbb{F}_q)$	simple?
2	S_3	No
3	A_4	No
4	A_5	Yes
5	A_5	Yes
7	$\cdot$	Yes
8	$\cdot$	Yes
9	A_6	Yes

Fact

Some other unexpected isomorphisms are

$$Sp_4(\mathbb{F}_2) \simeq S_6$$

$$SO_6(\mathbb{F}_2) \simeq SL_4(\mathbb{F}_2) \simeq A_8$$

§36 December 4, 2020

Example (Group "Prime Factorization")

In general, given a finite group G, if G has a normal subgroup N, then we can "factor" G into N and G/N; we can similarly "factor" N and G/N until all parts are simple. Eventually, you get a chain of subgroups:

$$1 = G_n \lhd G_{n-1} \lhd \cdots \lhd G_2 \lhd G_1 \lhd G_0 = G$$

with $G_{i+1} \lhd G_i$ and G_i/G_{i+1} simple for all i.

Theorem 77 (Classification of Finite Simple Groups)

Every finite simple group is one of the following:

- C_p for some prime p
- A_n for some $n \geq 5$
- $PSL_2(\mathbb{F}_q)$ for some prime power $q \geq 4$
- 15 other infinite families coming from matrix groups over finite fields
- 26 "sporadic" groups, the largest of which is called the Monster group

Definition (Exactness)

A sequence of homomorphisms, where A, B, C are groups and α, β are homomorphisms is a sequence

$$A \xrightarrow{\alpha} B \xrightarrow{\beta} C$$

is **exact** (at B) if $\text{im}(\alpha) = \ker(\beta)$.

Example

For instance, consider

$$\begin{aligned} \mathbb{R} \xrightarrow{\alpha} &\mathbb{C}^\times \xrightarrow{\beta} \mathbb{R}^\times \\ t \mapsto &e^{it} \\ z \mapsto &|z| \end{aligned}$$

is exact, since

$$\text{im}(\alpha) = \{e^{it} : t \in \mathbb{R}\} = \{z \in \mathbb{C}^\times : |z| = 1\} = \ker(\beta).$$

Definition (Exactness)

A sequence of homomorphisms

$$G_0 \xrightarrow{\alpha_1} G_1 \xrightarrow{\alpha_2} G_2 \xrightarrow{\alpha_3} \cdots \xrightarrow{\alpha_n} G_n$$

is **exact** if it is exact at each joint:

$$\text{im}(\alpha_i) = \ker(\alpha_{i+1}) \text{ for } i = 1, 2, \ldots, n-1$$

Example

The following are equivalent:

- $\{1\} \to A \xrightarrow{\phi} B$ is exact
- $\ker \phi = 1$
- ϕ is injective

Example

The following are equivalent:

- $B \xrightarrow{\psi} C \to 1$ is exact
- $\operatorname{im} \psi = C$
- ψ is surjective

Definition (Short Exact Sequence)

An exact sequence of the form

$$1 \to A \to B \to C \to 1$$

where 1 represents the trivial group, is called a **short exact sequence**.

Example

If $\phi : B \to C$ is a surjective homomorphism, then

$$1 \to \ker \phi \xrightarrow{\text{inclusion}} B \xrightarrow{\phi} C \to 1$$

is a short exact sequence.

At $\ker \phi$, the image coming in is 1, and the kernel going out is 1. At B, the image coming in is $\ker \phi$ and the kernel going out is $\ker \phi$. At C, the image coming in is C and the kernel going out is C.

Proposition

Every short exact sequence is isomorphic to the example above.

Proof. If

$$1 \to A \to B \xrightarrow{\phi} C \to 1$$

is exact, then $B \to C$ is surjective; call it ϕ. Then

$$\ker \phi = \operatorname{im}(A \to B) \simeq A$$

□

Example

If $N \lhd G$ (N is a normal subgroup of G), then

$$1 \to N \xrightarrow{\text{inclusion}} G \xrightarrow{\pi} G/N \to 1$$

where π is the canonical homomorphism is a short exact sequence.

Proof. At N, the incoming image and outgoing kernel are 1; at G, the incoming image and outgoing kernel are N; at G/N, the incoming image and outgoing kernel are G/N. □

Corollary

If $1 \to A \to B \to C \to 1$ is exact, then

$$|B| = |A|\,|C|\,.$$

§37 December 7, 2020

Fact

In the case of a short exact sequence

$$1 \to A \to B \to C \to 1,$$

then

- $A \to B$ is injective
- $B \to C$ is surjective
- A can be identified with a normal subgroup of B
- Given $A \lhd B$, we have $C \simeq B/A$
- Given $B \twoheadrightarrow C$, we have $A \simeq \ker(B \twoheadrightarrow C)$

Definition (Automorphism Group)

The automorphism group of G is

$$\operatorname{Aut} G := \{\text{automorphism of } G\}$$

which is a group under composition.

Example

For example, considering $C_5 = \langle g \rangle$ with $g^5 = 1$, for each integer r, there is a homomorphism

$$\begin{aligned} \alpha_r : C_5 &\to C_5 \\ g &\mapsto g^r \end{aligned}$$

if $5 \mid r$, then $g^r = 1$, so α_r sends C_5 to 1, so α_r is not an automorphism. On the other hand, if $5 \nmid r$, then g^r generates C_5, so α_r is an automorphism. In particular,

$$\alpha_r = \alpha_{r'} \iff g^r = g^{r'} \iff r \equiv r \pmod 5$$

and

$$\alpha_r \circ \alpha_5 = \alpha_{r5}$$

and there exists an isomorphism between $\operatorname{Aut} C_5 \simeq \mathbb{F}_5^\times$.

Example (Elementary Abelian Group)

We have that

$$\underbrace{C_p \times \cdots \times C_p}_{n} \simeq \mathbb{F}_p^n$$

and

$$\operatorname{Aut}(\underbrace{C_p \times \cdots \times C_p}_{n}) \simeq \operatorname{Aut}(\mathbb{F}_p^n) = GL_n(\mathbb{F}_p),$$

since GL_n is the automorphisms of the vector space.

Definition (Complement)

Suppose $N \lhd G$. Then, a **complement** of N in G is a subgroup $H \leq G$ such that

- $NH = G$
- $N \cap H = 1$

Example

For example, considering $\langle \rho \rangle = C_n \lhd D_n$, where ρ is a rotation by $2pi/n$ around $\vec{0}$ in $\mathbb{R}^2$, then a complement would be

$$H := \{1, r\}$$

where r is a reflection in the x-axis.

Proposition

If H is a complement of $N \lhd G$, then

$$\begin{aligned} N \times H &\to G \\ (x, h) &\mapsto xh \end{aligned}$$

is a bijection of sets.

Proof. This mapping is surjective by definition. Suppose (x_1, h_1) and (x_2, h_2) have the same image in G. Then

$$x_1 h_1 = x_2 h_2$$

so

$$x_2^{-1} x_1 = h_2 h_1^{-1}$$

is in $N \cap H$, and hence is 1. Thus, $x_1 = x_2$ and $h_1 = h_2$, so $(x_1, h_1) = (x_2, h_2)$. □

Example

We can check whether

$$N \times H \to G$$

is an isomorphism of groups, which is equivalent to checking if it is a homomorphism, since we know it's a bijection.

Given (x, h) and (x', h'), we could first multiply in $N \times H$ and then map, which would give $xx'hh'$. We could also first map into G and then multiply in G which gives $xhx'h'$. These two quantites are equal if and only if $x'h = hx'$. Thus, $N \times H \to G$ is an isomorphism of groups if and only if every element of N commutes with every element of H.

Fact

In general, for a complement H of $N \lhd G$, each $h \in H$ gives rise to

$$\begin{aligned} \mathrm{inn}_h : G &\to G \\ x &\mapsto hxh^{-1} \end{aligned}$$

which restricts to an automorphism

$$\phi_h : N \to N$$

since N is normal. This yields a homomorphism

$$\begin{aligned} H &\xrightarrow{\phi} \operatorname{Aut} N \\ h &\mapsto \phi_n \end{aligned}$$

where we can call ϕ an action of H on the group N.

By definition, if $h \in H$ and $x \in N$, then

$$hxh^{-1} = \phi_h(x),$$

another element of N. Thus

$$hx = \phi_h(x)h.$$

This determines the multiplication on all of G:

$$(xh)(x'h') = x(hx')h' = \underbrace{x\phi_h(x')}_{\in N}\underbrace{hh'}_{\in H}.$$

Example

We can consider the general construction. Given groups N and H and a homomorphism

$$\begin{aligned} H &\to \operatorname{Aut} N \\ h &\mapsto \phi_h \end{aligned}$$

we can define the semidirect product $N \rtimes_\phi H$ to be the set $N \times H$ with the binary operation

$$(x, h) \cdot (x', h') := (x\phi_h(x'), hh').$$

Proposition

$N \rtimes_\phi H$ is a group known as the semidirect product of N and H with respect to homomorphism ϕ.

We can prove this simply by checking the group axioms. We can also consider the two subgroups:

$$\{(x, 1) : x \in N\}$$

which is a copy of N in $N \rtimes H$ and

$$\{(1, h) : h \in H\}$$

which is a copy of H in $N \rtimes H$. In particular, N is a normal subgroup and we can get a short exact sequence

$$1 \to N \to N \rtimes H \to H \to 1.$$

§38 December 9, 2020

Proposition

If H is a complement of $N \lhd G$, then

- H acts on N by conjugation, via

$$\begin{aligned} \phi : H &\to \operatorname{Aut} N \\ h &\mapsto (x \mapsto hxh^{-1}). \end{aligned}$$

- $G \simeq N \rtimes_\phi H$.

Example

Every group of order pq for primes p, q is a semidirect product. For instance, the nonabelian group of order 21 is isomorphic to

$$C_7 \rtimes_\phi C_3$$

where $\phi : C_3 \to \operatorname{Aut} C_7 = \mathbb{F}_7^\times$, mapping $1 \mapsto 1, g \mapsto 2, g^2 \mapsto 4$.

Example

Given V and W, $\mathbb{R}$-vector spaces, and given $v \in V$ and $w \in W$, can we "multiply" to get some $v \otimes w$? In which vector space would it lie?

Supposing we have basis vectors $(e_1, e_2, e_3) \in \mathbb{R}^3$ and $(f_1, f_2) \in \mathbb{R}^2$, we can take the tensor product by forming the products

$$e_1 \otimes f_1, e_2 \otimes f_1, e_3 \otimes f_1, e_1 \otimes f_2, e_2 \otimes f_2, e_3 \otimes f_2$$

which are a basis of $\mathbb{R}^6$. In particular, there is a unique way to textend the $\otimes$ on basis vectors to all vectors, to define $v \otimes w$ for all $v \in \mathbb{R}^3$, $w \in \mathbb{R}^2$ in such a way that

$$\begin{aligned} \mathbb{R}^3 \times \mathbb{R}^2 &\to \mathbb{R}^3 \otimes \mathbb{R}^2 \\ (v, w) &\mapsto v \otimes w \end{aligned}$$

is a bilinear function.

Definition (Tensor Product)

We can let R be the subapce of G spanned by

$$\begin{cases} (v_1 + v_2, w) - (v_1, w) - (v_2, w) \\ (cv, w) - c(v, w) \\ (v, w_1 + w_2) - (v, w_1) - (v, w_2) \\ (v, cw) - c(v, w) \end{cases}$$

for all $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$, and $c \in \mathbb{R}$. Then, we can define

$$V \otimes W := G/R$$

such that

$$v \otimes w := \text{the image of } (v, w) \in G \text{ under } G \twoheadrightarrow G/R.$$

Then, for example,

$$(v_1 + v_2, w) - (v_1, w) - (v_2, w)$$

is in R, so

$$(v_1 + v_2) \otimes w - v_1 \otimes w - v_2 \otimes w = 0$$

is the image in G/R, meaning

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w.$$

We can similarly verify the other properties of bilinearity. Notice that the same construction works over any field F. If V and W are F-vector spaces, then we can form

$$V \otimes_F W$$

and a function

$$\begin{aligned} V \times W &\to V \otimes W \\ (v, w) &\mapsto v \otimes w \end{aligned}$$

Example

For example,

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} = \text{4-dimensional } \mathbb{R}\text{-vector space}$$

and

$$\mathbb{C} \otimes_{\mathbb{C}} \mathbb{C} = \text{1-dimensional } \mathbb{C}\text{-vector space}$$

Fact

$\mathbb{R}$ is a $\mathbb{Q}$-vector space (under addition and scalar multiplication by rational numbers) of infinite dimension. For instance,

$$\sqrt{2}, \sqrt{3}, \sqrt{5}, \ldots \text{ are } \mathbb{Q}\text{-linearly dependent.}$$

$\mathbb{Q}_\pi$ is a 1-dimensional $\mathbb{Q}$-subspace of $\mathbb{R}$. Then, $\mathbb{R}/\mathbb{Q}\pi$ is another infinite-dimensional vector space.

Definition (Scissors-Congruent)

Suppose you have a polyhedra P, which you cut with planes and reattach the pieces to make P'. Then, P and P' are **scissors-congruent**. These scissors-congruent polyhedra preserve two properties:

- P and P' have the same volume
- the Dehn invariant of P, $\sum_{e \in P} l_e \otimes \theta_e \in \mathbb{R} \otimes_{\mathbb{Q}} \mathbb{R}/\mathbb{Q}\pi$

where l_e is the length of edge e and θ_e is the dihedral angle (angle between faces) along edge e.

Theorem 78 (Dehn-Sydler)

Polyhedra P and P' are scissors-congruent if and only if the volumes of P and P' are equal, and the Dehn inavariants of P and P' are equal.

Example

This allows us to conclude that a regular tetrahedron cannot be cut up and reassembled into a cube!