18.701 — Algebra 1

Class by Davesh Maulik

Notes by Sanjana Das

Fall 2021

Notes for the MIT class **18.701** (Algebra I), taught by Davesh Maulik. All errors are my responsibility.

Contents

§1 Groups

Before we define groups, we'll introduce a useful example.

§1.1 The General Linear Group

Definition 1.1. The *general linear group*, denoted $\mathrm{GL}_n(\mathbb{R})$, is the set of invertible $n \times n$ matrices with real coefficients.

Recall that a matrix A is defined to be invertible if there exists another $n \times n$ matrix, denoted A^{-1}, such that $AA^{-1} = A^{-1}A = I$. A matrix is invertible if and only if $\det(A) \neq 0$.

This set of invertible matrices has an interesting property — if we multiply two invertible matrices A and B, then we get another, since AB has inverse $B^{-1}A^{-1}$. This multiplication is *not* commutative — in general, AB does not equal BA. But it *is* associative — we always have $A(BC) = (AB)C$.

Remark 1.2. It's generally useful to think of matrices as *operations* on some space. Given a matrix $A \in \mathrm{GL}_n(\mathbb{R})$, we can define a function $\mathbb{R}^n \to \mathbb{R}^n$ sending $v \mapsto Av$. This function must be linear; and conversely, given any linear function, we can recover its corresponding matrix.

In this interpretation, all these facts become much more intuitive — multiplying two matrices corresponds to composing the functions, and function composition is associative but not commutative. (This is where the definition of matrix multiplication comes from — multiplication is *defined* to match what happens when we compose the two maps.)

With this example in mind, we can now define groups in general.

§1.2 What Is a Group?

Definition 1.3. A *group* is a set G with a composition (or product) operation $G \times G \to G$, denoted by $(a, b) \mapsto a \cdot b$ (or ab), that satisfies the following axioms:

(1) *Identity* — there exists $e \in G$ such that $ae = ea = a$ for all $a \in G$.

(2) *Inverses* — for each $a \in G$, there exists an element $b \in G$ such that $ab = ba = e$. We call b the *inverse* of a, and write $b = a^{-1}$.

(3) *Associativity* — we have $(ab)c = a(bc)$ for all $a, b, c \in G$.

One observation we can immediately make from these axioms is that in (1) and (2), the element must be unique. In (1), if both e and e' are identity elements, then ee' must equal both e' and e, so $e' = e$. Similarly in (2), if b and b' are both inverses of a, then bab' must equal both b' and b, so $b' = b$.

In this definition, we've only defined the product of two elements. But thanks to associativity, we can talk about products of many elements, such as $g_1g_2 \cdots g_n$, without having to specify in what order we pair up the elements when calculating. In particular, for nonnegative integers n we write g^n to denote the product of n copies of g; similarly for negative n we write g^n to denote the product of $-n$ copies of g^{-1}.

Example 1.4

Some examples of groups are:

(1) $\mathrm{GL}_n(\mathbb{R})$ under matrix multiplication — the identity is I, and the inverse of A is the matrix A^{-1}.

(2) $\mathbb{Z}$ under addition — the identity is 0, and the inverse of a is $-a$.

(3) $\mathbb{C}^\times$ (complex numbers except for 0) under multiplication — the identity is 1, and the inverse of x is $1/x$.

As we've seen already, the composition law doesn't have to be commutative. But if it is, we get additional structure, so such groups have a name:

Definition 1.5. A group G is *abelian* if $ab = ba$ for all $a, b \in G$.

Example 1.6

The groups $\mathbb{Z}$ and $\mathbb{C}^\times$ are both abelian, while $\mathrm{GL}_n(\mathbb{R})$ is not (for $n \geq 2$).

§1.3 Permutation Groups

Another central example of a group is the permutation group.

Definition 1.7. Given a set S, a *permutation* of S is a bijection $p\colon S \to S$.

The set of all permutations of S, denoted $\mathrm{Perm}(S)$, is a group under function composition — for two permutations p and q, we define their product $q \cdot p$ as the permutation $q \cdot p(x) = q(p(x))$ for each x. The identity is the identity permutation, for which $e(x) = x$ for all x. Inverses exist because bijections have inverses — for a permutation $p \in \mathrm{Perm}(S)$, its inverse p^{-1} is the permutation where for each x, we define $p^{-1}(x)$ to be the unique y with $p(y) = x$.

Note that we can think of $\mathrm{Perm}(S)$ as the group of *symmetries* on S — in much later classes, we will explore how to think of groups via symmetry in more detail.

Definition 1.8. The group of permutations of $\{1, 2, \ldots, n\}$ is called the *symmetric group* S_n.

Unlike all our previous examples of groups, S_n is finite.

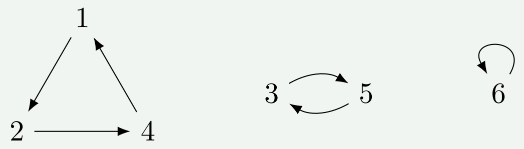
Definition 1.9. The *order* of a group is its number of elements.

So the order of S_n is $n!$, since there are $n!$ permutations of $\{1, 2, \ldots, n\}$.

It's often useful to describe permutations using *cycle notation* — we can draw arrows $i \mapsto p(i)$ for each element i, and write down a permutation by writing down all its cycles.

Example 1.10

The permutation p sending 1, 2, 3, 4, 5, 6 to 2, 4, 5, 1, 3, 6, respectively, would have cycle notation (124)(35)(6). We may also omit cycles of length 1, and write (124)(35).



Note that cycle notation can be thought of as taking the *composition* of disjoint cycles — for example, the permutation (135)(246) is $(135) \cdot (246)$.

To find the inverse of a permutation p given in cycle notation, we can simply reverse each cycle — in this example, $p^{-1} = (421)(53) = (142)(35)$.

We can also compute compositions using cycle notation:

Example 1.11

Let $p = (124)(35)$ and $q = (135)(246)$. Then in order to compute $q \cdot p$, we first find where p sends i, and then where q sends $p(i)$ — this gives $q \cdot p = (143)(26)$.

Finally, another interesting operation we can perform is *conjugation* — given two permutations p and q, we can calculate $p^{-1} \cdot q \cdot p$, which is called the conjugate of q by p.

Example 1.12

If $p = (124)(35)$ and $q = (135)(246)$, then $p^{-1} \cdot q \cdot p = (126)(345)$.

In an abelian group, the conjugate of q by p will always be q itself; but here S_6 is not abelian, so we got a new permutation. We will see conjugation in more detail later; but it's not a coincidence that $p^{-1} \cdot q \cdot p$ here has the same "shape" of its cycles as q does.

Finally, we'll explicitly describe the symmetric group for a small value of n.

Example 1.13

The group S_3 contains e, the element $x = (123)$, and the element $y = (12)$. The remaining elements can be described as $x^2 = (132)$, $xy = (13)$, and $x^2y = (23)$.

Our elements x and y satisfy the relations $x^3 = e$, $y^2 = e$, and $yx = x^2y$. These together are enough to reduce *any* combination of x and y to one of the six forms listed — for example,

$$xyx^{-1}y = xyxxy = xx^2yxy = yxy = x^2y^2 = x^2.$$

§1.4 Subgroups

Given a group, we can also look at smaller groups which sit inside it.

Definition 1.14. Given a group $(G, \cdot)$, a subset $H \subset G$ is a *subgroup* of G if it satisfies the following conditions:

(1) *Closure* — if $h_1, h_2 \in H$, then $h_1 \cdot h_2 \in H$ as well.

(2) *Identity* — $e \in H$, where e is the identity element of G.

(3) *Inverses* — for each $h \in H$, we have $h^{-1} \in H$.

Equivalently, H is a subgroup of G if it is also a group, under the same operation.

Example 1.15

Some examples of subgroups are:

(1) $(\mathbb{Z}, +)$ is a subgroup of $(\mathbb{Q}, +)$.

(2) $\{e, (123), (132)\}$ is a subgroup of S_3.

(3) As a nonexample, $\mathbb{Z}_{\geq 0}$ is *not* a subgroup of $\mathbb{Z}$, since it is not closed under taking inverses.

(4) The *special linear group* $\mathrm{SL}_n(\mathbb{R})$, consisting of matrices with determinant 1, is a subgroup of $\mathrm{GL}_n(\mathbb{R})$ — it is closed under both multiplication and taking inverses because determinants are multiplicative.

Notation 1.16. The notation $H \leq G$ can be used to denote that H is a subgroup of G.

§1.4.1 Subgroups of $\mathbb{Z}$

It turns out that subgroups of the integers are quite easy to describe.

Theorem 1.17

The subgroups of $(\mathbb{Z}, +)$ are exactly $\{0\}$ and $n\mathbb{Z}$ for positive integers n.

Proof. It's easy to check that all such sets are subgroups of $\mathbb{Z}$; now we'll show that any subgroup of $\mathbb{Z}$ must be one of these.

Let $S \subset \mathbb{Z}$ be a subgroup. We must have $0 \in S$; if there are no other elements, we're done. Otherwise, let a be the smallest positive element (which exists because if x is in S, so is $-x$). Now we claim that $S = a\mathbb{Z}$ — first, by closure we must have $a\mathbb{Z} \subset S$. Now assume for contradiction there is some $b \in S$ which is not a multiple of a. Then by closure, $b - ka$ must be in S for all integers k. In particular, the remainder when b is divided by a must also be in S. But this remainder is strictly between 0 and a (since $a \nmid b$), contradiction. So all elements in S are multiples of a, and therefore $S = a\mathbb{Z}$. □

Corollary 1.18

Given integers a and b, let $S = \{ai + bj \mid i, j \in \mathbb{Z}\}$. Then $S = d\mathbb{Z}$ for some positive integer d.

Proof. We can check that S is a (nonzero) subgroup of $\mathbb{Z}$; but all subgroups of $\mathbb{Z}$ are either $\{0\}$ or of the form $d\mathbb{Z}$ for a positive integer d. □

This leads to an important result in elementary number theory:

Theorem 1.19 (Bezout's Theorem)

Given integers a and b, there exist integers r and s for which

$$ar + bs = \gcd(a, b).$$

Proof. We'll show that in the above corollary, we must have $d = \pm\gcd(a, b)$. This suffices because then $\gcd(a, b)$ is in S.

First, all numbers $ai + bj$ are multiples of $\gcd(a, b)$, so since d is in S and therefore can be written as $ai + bj$ for some i and j, then $\gcd(a, b)$ must divide d. On the other hand, since a and b are in S, and S consists of exactly the multiples of d, then d must divide a and b, and therefore must divide $\gcd(a, b)$ as well. So since d and $\gcd(a, b)$ both divide each other, we have $d = \pm\gcd(a, b)$. □

Remark 1.20. Bezout's Theorem can be extended to multiple integers, instead of just two — in general, given any integers $a_1, \ldots, a_n$, there exist integers $r_1, \ldots, r_n$ such that

$$a_1r_1 + \cdots + a_nr_n = \gcd(a_1, \ldots, a_n).$$

§1.5 Cyclic Groups

One of the simplest examples of a group is a *cyclic* group.

Definition 1.21. Given an element g of a group G, the *cyclic group* generated by g, denoted $\langle g\rangle$, is the smallest subgroup of G containing g.

Then we have $\langle g\rangle = \{\ldots, g^{-2}, g^{-1}, e, g^1, g^2, \ldots\}$ — if a subgroup of G contains g then by closure it muts contain all powers of g, while this set is really a valid group.

Example 1.22

Some examples of cyclic groups are:

(1) In any group, we have $\langle e\rangle = \{e\}$.

(2) In S_3, we have $\langle(123)\rangle = \{e, (123), (132)\}$.

(3) In $\mathbb{C}^\times$, we have $\langle 2\rangle = \{\ldots, \frac{1}{4}, \frac{1}{2}, 1, 2, 4, \ldots\}$ and $\langle i\rangle = \{1, i, -1, -i\}$.

Question 1.23. What do cyclic groups look like in general?

To answer this, let S be the set of integers n for which $g^n = e$.

Theorem 1.24

Either $S = \{0\}$, in which case $\langle g\rangle$ is infinite and all the powers g^i are distinct; or $S = d\mathbb{Z}$, in which case $\langle g\rangle$ is finite and contains exactly d elements — more precisely, $\langle g\rangle = \{e, g, g^2, \ldots, g^{d-1}\}$.

Proof. First we claim that S is a subgroup of $\mathbb{Z}$:

- 0 is in S by definition, since $g^0 = e$.

- If $g^a = g^b = e$, then $g^{a+b} = g^a \cdot g^b = e$ as well. So if a and b are in S, so is $a + b$.
- If $g^a = e$, then $g^{-a} = e^{-1} = e$ as well. So if a is in S, so is $-a$.

But the only subgroups of $\mathbb{Z}$ are $\{0\}$ and $d\mathbb{Z}$ for positive integers d, so S must be one of these.

Now note that for two integers a and b, we have $g^a = g^b$ if and only if $g^{a-b} = e$. So if $S = \{0\}$, this implies $a = b$; therefore all powers of g are distinct, and $\langle g \rangle$ is infinite. Meanwhile, if $S = d\mathbb{Z}$, then this means $g^a = g^b$ if and only if $a \equiv b \pmod{d}$. So every element of $\langle g \rangle$ is in $\{e, g, g^2, \dots, g^{d-1}\}$, and these d powers are all distinct. □

This gives the following definition:

> **Definition 1.25.** The *order* of an element g of a group is defined as $\operatorname{ord}(g) = \#\langle g \rangle$.

In other words, $\operatorname{ord}(g)$ is the smallest positive integer d for which $g^d = e$ if such an integer exists, and infinite otherwise.

§1.6 Generators

Given a group G, we've seen what happens when we look at the smallest subgroup of G containing *one* given element g. But we can also look at the smallest subgroup containing *multiple* elements:

> **Definition 1.26.** Given a group G and a subset $T \subset G$, the subgroup generated by T is the smallest subgroup of G that contains T.

For the same reason as $\langle g \rangle$ consists of all powers of g, in general $\langle T \rangle$ consists of all products of powers of elements in T — more explicitly, we have

$$\langle T \rangle = \{t_1^{a_1} t_2^{a_2} \cdots t_n^{a_n} \mid t_i \in T, a_i \in \mathbb{Z} \text{ for all } i\}.$$

(Note that the t_i do not have to be distinct.)

> **Definition 1.27.** Given a group G and a subset $T \subset G$, if $\langle T \rangle = G$ then we say T generates G.

Example 1.28

As we saw in Example 1.13, S_3 is generated by $\{(123), (12)\}$.

Example 1.29

The group $\mathrm{GL}_n(\mathbb{R})$ is generated by the *elementary matrices* — matrices corresponding to the elementary row operations.

§1.7 Homomorphisms

When we've defined a structure — here, a group — we can then ask how two such structures can relate to each other, by looking at the maps between them.

> **Definition 1.30.** Given groups G and H, a *homomorphism* from G to H is a map $f\colon G \to H$ such that for all $x, y \in G$,
> $$f(x \cdot y) = f(x) \cdot f(y).$$

Note that the multiplication $x \cdot y$ occurs in G, while the multiplication $f(x) \cdot f(y)$ occurs in H — these may be different operations.

Essentially, a homomorphism is a map between two groups which is compatible with their group structures. There are other conditions it would make sense to include as well to describe compatibility with the group structure — we would want f to also preserve the identity and inverses. However, it turns out it's not necessary to state these conditions in the definition, because they follow from the given one!

Proposition 1.31

For any homomorphism $f\colon G \to H$, we must have $f(e_G) = e_H$, and $f(x^{-1}) = f(x)^{-1}$ for all $x \in G$.

Proof. To prove the first property, for any $x \in G$ we have

$$f(x) = f(e_G \cdot x) = f(e_G) \cdot f(x).$$

Multiplying by $f(x)^{-1}$ on both sides (on the right), we get that $f(e_G) = e_H$.

Now to prove the second, for any x we have

$$e_H = f(e_G) = f(x \cdot x^{-1}) = f(x) \cdot f(x^{-1}),$$

so $f(x^{-1})$ must be the inverse of x. □

§1.7.1 Examples

Example 1.32

The map $\det\colon \mathrm{GL}_n(\mathbb{R}) \to (\mathbb{R}^\times, \times)$ is a homomorphism, since $\det(A)\det(B) = \det(AB)$ for any two matrices A and B.

Example 1.33

The map $\exp\colon (\mathbb{C}, +) \to (\mathbb{C}^\times, \times)$, which is defined as $z \mapsto e^z$, is a homomorphism, since $e^{a+b} = e^a e^b$ for any $a, b \in \mathbb{C}$.

Another important homomorphism, from the group S_n, is the *sign* of a permutation.

First, let $\vec{e_i}$ denote the column vector with a 1 in the ith coordinate and a 0 everywhere else. Then for each permutation $p \in S_n$, we can associate to it a *permutation matrix* A_p, defined such that $A_p(\vec{e_i}) = \overrightarrow{e_{p(i)}}$ for all i. In other words, A_p corresponds to the linear map which permutes the basis vectors $\vec{e_1}, \ldots, \vec{e_n}$ in the same way as p permutes $\{1, \ldots, n\}$.

Example 1.34

The permutation matrix associated to $p = (123) \in S_3$ is

$$A_p = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Proposition 1.35

The function $S_n \to \mathrm{GL}_n(\mathbb{R})$ sending $p \mapsto A_p$ for each permutation p is a homomorphism.

Proof. It's enough to check that $A_{pq} = A_p A_q$ for all permutations p and q. First, by definition we have $A_{pq}(\overrightarrow{e_i}) = \overrightarrow{e_{pq(i)}}$ for each basis vector $\overrightarrow{e_i}$. On the other hand,

$$A_p A_q(\overrightarrow{e_i}) = A_p(\overrightarrow{e_{q(i)}}) = \overrightarrow{e_{p(q(i))}} = \overrightarrow{e_{pq(i)}}.$$

So then A_{pq} and $A_p A_q$ are the same map. □

But we also have a homomorphism $\mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$, the determinant. So we can compose these two homomorphisms to get a new one:

Definition 1.36. The *sign function* $\operatorname{sgn}\colon S_n \to \mathbb{R}^\times$ is defined as $\operatorname{sgn}(p) = \det(A_p)$ for each $p \in S_n$.

Proposition 1.37

We have $\operatorname{sgn}(p) = \pm 1$ for any permutation p.

Proof. Every permutation can be written as a product of transpositions $p = \tau_1 \tau_2 \cdots \tau_r$ — for example, by using any sorting algorithm that only involves making swaps. But for any transposition τ, the matrix A_τ can be obtained by swapping two rows of the identity matrix, which means $\operatorname{sgn}(\tau) = -1$ (since swapping rows multiplies the determinant by -1). So then since sgn is a homomorphism, in general we have

$$\operatorname{sgn}(p) = \operatorname{sgn}(\tau_1)\operatorname{sgn}(\tau_2)\cdots\operatorname{sgn}(\tau_r) = (-1)^r.$$ □

Note that this also implies that when we write p as a product of r transpositions, the parity of r must be fixed! Often, a permutation is called *even* or *odd* depending on the parity of r.

Example 1.38

In S_3, the permutations e, (123), and (132) have sign 1, and (12), (13), and (23) have sign -1.

Finally, here is another example of a homomorphism, which in fact we've seen already:

Example 1.39

For any x in a group G, there is a homomorphism $f_x\colon \mathbb{Z} \to G$ sending $n \mapsto x^n$. This is a homomorphism because $x^{a+b} = x^a \cdot x^b$ for any integers a and b.

We secretly used this homomorphism when studying the cyclic group generated by x. As in that example, homomorphisms are useful because they can be used to study complicated groups in terms of simpler ones.

§1.7.2 Image and Kernel

In all definitions here, we assume that f is a homomorphism $f\colon G \to H$.

Definition 1.40. The *image* of f, denoted $\operatorname{im}(f)$, is the set of elements $y \in H$ such that $y = f(x)$ for some $x \in G$.

Example 1.41

By definition, the image of the homomorphism f_x in Example 1.39 is $\langle x \rangle$.

Proposition 1.42

The image of f is a subgroup of H.

Proof. To show that $\text{im}(f)$ is closed under multiplication, suppose y and y' are in the image of f, so we have $y = f(x)$ and $y' = f(x')$ for some x and x'. Then

$$yy' = f(x)f(x') = f(xx'),$$

so yy' is also in the image of f. It's possible to check the other conditions — that it's closed under taking inverses, and it contains the identity — in a similar way. □

Definition 1.43. The *kernel* of f, denoted $\ker(f)$, is the set of elements $x \in G$ for which $f(x) = e_H$.

Proposition 1.44

The kernel of f is a subgroup of G.

Proof. First, if x and y are both in $\ker(f)$, then $f(x) = f(y) = e_H$, so

$$f(xy) = f(x)f(y) = e_H$$

as well, and therefore xy is also in $\ker(f)$. So $\ker(f)$ is closed under multiplication.

We already showed that any homomorphism must satisfy $f(e_G) = e_H$, so $\ker(f)$ contains the identity of G.

Finally, if x is in $\ker(f)$, then $f(x) = e_H$, so

$$f(x^{-1}) = f(x)^{-1} = e^H$$

as well. So $\ker(f)$ is closed under taking inverses as well. Therefore $\ker(f)$ is a subgroup of G. □

Example 1.45

For the homomorphism $f_x\colon \mathbb{Z} \to G$ defined as $n \mapsto x^n$ (for a fixed element $x \in G$), the kernel of f_x is precisely the set of n such that $x^n = e_G$. This is exactly the set S we used in order to describe $\langle x \rangle$ — in particular, we used the fact that it's a subgroup of $\mathbb{Z}$. More explicitly, this kernel is $d\mathbb{Z}$ if $d = \text{ord}(x)$ is finite, and $\{0\}$ if $\text{ord}(x)$ is infinite.

Example 1.46

The images and kernels for the other homomorphisms described in the previous section are the following:

(1) For the map $\det\colon \text{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$, the image is $\mathbb{R}^\times$, and the kernel is $\text{SL}_n(\mathbb{R})$, which denotes the group of matrices with determinant 1.

(2) For the map $\exp\colon \mathbb{C} \to \mathbb{C}^\times$, the image is $\mathbb{C}^\times$ and the kernel is $2\pi i\mathbb{Z}$ (the cyclic group generated by $2\pi i$).

(3) For the map $S_n \to \text{GL}_n(\mathbb{R})$ defined as $p \mapsto A_p$, the image is the set of all permutation matrices, and the kernel is the identity permutation.

(4) For the map $\text{sgn}\colon S_n \to \mathbb{R}^\times$, the image is $\{\pm 1\}$ and the kernel is the set of permutations with sign $+1$. The kernel of sgn is called the *alternating group* and denoted by A_n. For example, we have $A_3 = \{e, (123), (132)\}$.

In some sense, the kernel measures the failure of f to be injective — if the kernel is trivial, then f is injective.

§1.7.3 Isomorphisms

Definition 1.47. A bijective homomorphism is called an *isomorphism.*

Recall that a function $f\colon G \to H$ is bijective if it's both *surjective*, meaning that all of H is in its image, and *injective*, meaning that it sends any two distinct elements of G to distinct elements of H.

Example 1.48

The map $\exp\colon (\mathbb{R}, +) \to (\mathbb{R}_{>0}, \times)$ defined as $t \mapsto e^t$ is an isomorphism.

Claim 1.49 — Given an isomorphism $f\colon G \to H$, its inverse $f^{-1}\colon H \to G$ is also an isomorphism.

Proof. First since f is bijective, it has an inverse f^{-1}, which is bijective as well. So it suffices to check that

$$f^{-1}(xy) = f^{-1}(x)f^{-1}(y)$$

for all $x, y \in H$. To check this, we can take f of both sides — we have

$$f(f^{-1}(xy)) = xy = f(f^{-1}(x))f(f^{-1}(y)) = f(f^{-1}(x)f^{-1}(y)).$$

But since f is injective, this means $f^{-1}(xy) = f^{-1}(x)f^{-1}(y)$. □

If we have an isomorphism between two groups, then they're basically the same — anything that can be said about one can also be said about the other, by just renaming elements according to the isomorphism. So understanding the group is essentially the same as understanding a group isomorphic to it.

Example 1.50

For an element $g \in G$ with finite order d, the map f_x gives an isomorphism between $\mathbb{Z}/d\mathbb{Z}$ and $\langle g \rangle$, by sending a residue n mod d to the element g^n. Meanwhile, for an element $g \in G$ with infinite order, the map f_x is an isomorphism between $\mathbb{Z}$ and $\langle g \rangle$.

In order to check that this example (or more generally, any map) is an isomorphism, we'd check that it's a bijection and that it's compatible with the group operations.

Definition 1.51. An isomorphism from a group G to itself is called an *automorphism.*

Example 1.52

For any group, the identity map $g \mapsto g$ is an automorphism. But there are often more interesting automorphisms as well:

(1) The map $\mathbb{Z} \to \mathbb{Z}$ sending $n \mapsto -n$ is an automorphism. In fact, this map and the identity are the only automorphisms of $\mathbb{Z}$.

(2) The map on $\mathrm{GL}_n(\mathbb{R})$ sending $A \mapsto (A^\intercal)^{-1}$ is an automorphism.

In fact, there's a general construction that usually produces interesting automorphisms:

Example 1.53

For any group G and any element $a \in G$, the map $\varphi_a \colon G \to G$ sending $x \mapsto axa^{-1}$ is an automorphism. This is called *conjugation* by a, and these automorphisms are called *inner automorphisms.*

Proof. First we'll check that φ_a is a homomorphism: for any x and y, we have

$$\varphi_a(x)\varphi_a(y) = axa^{-1} \cdot aya^{-1} = axya^{-1} = \varphi_a(xy),$$

as desired. Meanwhile to check that φ_a is a bijection, note that its inverse is the map $\varphi_{a^{-1}} \colon x \mapsto a^{-1}xa$, since for any x we have $a(a^{-1}xa)a^{-1} = x$. □

This construction always produces *some* automorphisms but depending on the choice of a, this may or may not be an interesting one. In particular, if G is abelian, then φ_a is always just the identity map. However, when G is not abelian, conjugation can be interesting.

Remark 1.54. Given any group, its automorphisms *themselves* form a group, under function composition. This new group can be an interesting object to study.

§1.8 Cosets

Question 1.55. Given a homomorphism $f \colon G \to G'$, when do we have $f(a) = f(b)$?

We have $f(a) = f(b)$ if and only if $f(a)^{-1}f(b) = e$, or equivalently if and only if $f(a^{-1}b) = e$. This occurs exactly when $a^{-1}b$ is in the kernel of f, or in other words, when

$$b \in a\ker(f) = \{ax \mid x \in \ker(f)\}.$$

This motivates us to study what such sets look like.

Definition 1.56. Given a subgroup $H \leq G$, a *left coset* of H is a subset of G of the form $aH = \{ax \mid x \in H\}$.

Example 1.57

In the group $S_3 = \langle x, y \rangle$ where $x = (123)$ and $y = (12)$, find the left cosets of $\langle y \rangle = \{e, y\}$.

Solution. Let $H = \{e, y\}$. Taking a to be e, x, and x^2, we get the cosets $H = \{e, y\}$, $xH = \{x, xy\}$, and $x^2H = \{x^2, x^2y\}$. Now the remaining values of a give $yH = \{y, e\}$, which is the same as H; $xyH = \{xy, x\}$, which is the same as xH; and $x^2yH = \{x^2y, x^2\}$, which is the same as x^2H.

So the cosets of H are $\{e, y\}$, $\{x, xy\}$, and $\{x^2, x^2y\}$. Note that although there were six possible values of a to shift by, some produce the same coset — so there's only three different cosets. □

Example 1.58

In the group $\mathbb{Z}$, find the cosets of the subgroup $2\mathbb{Z}$.

Solution. If we shift by 0, then we get the even integers $2\mathbb{Z}$, and if we shift by 1, then we get the odd integers $2\mathbb{Z}+1$. These are the only two cosets — shifting by *any* even number produces $2\mathbb{Z}$, and shifting by any odd number produces $2\mathbb{Z}+1$. So here there are only two cosets. □

Proposition 1.59

All cosets of H have the same order as H.

Proof. The function $x \mapsto ax$ is a bijection from H to aH, since it has an inverse $x \mapsto a^{-1}x$. □

Proposition 1.60

The cosets of H form a partition of G.

A *partition* of a set S is a subdivision of S into disjoint subsets — so these subsets don't overlap, and together they contain all elements of S.

In order to prove this, we'll first prove the following lemma:

Lemma 1.61

Given a coset C of H, then for any element $b \in C$, we have $C = bH$.

Proof. Suppose $C = aH$ for some a. Then we have $b = ah$ for some $h \in H$, since b is in aH. This means

$$bH = \{bh' \mid h' \in H\} = \{ahh' \mid h' \in H\}.$$

But since h and h' are both in H, so is hh', and therefore $bH \subset aH$. We can show $aH \subset bH$ similarly, by writing $a = bh^{-1}$. So then aH and bH must be the same coset. □

Using this, we can now prove the proposition:

Proof of Proposition 1.60. First, every element is in a coset — since $e \in H$, we have $x \in xH$ for all elements x. Now to show that the cosets are disjoint, suppose C and C' are two cosets with nonempty intersection. Then if both cosets contain y, the above lemma implies that both are yH, so they are actually the same coset. So distinct cosets of H don't overlap. □

Definition 1.62. The *index* of a subgroup $H \leq G$, denoted $[G : H]$, is the number of left cosets of H.

Theorem 1.63

We have $\#G = [G : H] \cdot \#H$.

Proof. The cosets of H form a partition of G. But there are $[G : H]$ such cosets and each has size $\#H$, which gives the desired result. □

As a corollary, we get Lagrange's Theorem:

Theorem 1.64 (Lagrange's Theorem)

For any subgroup $H \leq G$, we have that $\#H$ divides $\#G$.

Corollary 1.65

If $\#G$ is prime, then G is cyclic.

Proof. Let $\#G = p$. For any element $x \in G$, we have that $\langle x \rangle$ is a subgroup of G. Now pick any $x \in G$ other than the identity, so $\langle x \rangle$ does not have order 1 (as it contains both e and x). However, its order must divide p, so it must *equal* p. This means $\langle x \rangle$ is the entire group, and therefore G is cyclic. Furthermore, this proof implies G is generated by *any* one of its non-identity elements. □

This means for every prime p, there's a *unique* group of order p up to isomorphism — every such group is isomorphic to the cyclic group $\mathbb{Z}/p\mathbb{Z}$.

We can generalize the argument used here:

Corollary 1.66

For any element x of a group G, $\text{ord}(x)$ must divide $\#G$.

Proof. The order of x is the size of $\langle x \rangle$, and since $\langle x \rangle$ is a subgroup of G, its size must divide that of G. □

Example 1.67

What are the possible groups of order 4?

Solution. The group must contain the identity, and every other element must have order 2 or 4. First, if there is some element x of order 4, then the group must be $\langle x \rangle \cong \mathbb{Z}/4\mathbb{Z}$.

Now assume that every element other than the identity has order 2. Take some element x; then since taking powers of x doesn't give any new element, there must exist some other element y. Now our group contains the elements e, x, y, and xy. Note that xy has to be a new element — if $xy = e$ then since $x^2 = e$ as well we would have $x = y$, while if xy were equal to x or y then the other one would equal e.

So the group consists of the four elements $\{e, x, y, xy\}$. But the same reasoning shows that the group consists of the elements $\{e, x, y, yx\}$, so we must have $xy = yx$. This means the group is exactly the elements $\{e, x, y, xy\}$ with the relations $x^2 = y^2 = e$ and $yx = xy$, which is enough to completely describe the group.

This group is abelian but not cyclic. In fact, it's isomorphic to the group of matrices

$$\begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix} \leq \text{GL}_2(\mathbb{R}).$$

So any group of size 4 is either cyclic or isomorphic to this second group. □

With this, we can return to our original question about when we have $f(a) = f(b)$ given a homomorphism $f\colon G \to G'$. For each $y \in G'$, we can consider its pre-image

$$\{f^{-1}(y) = \{x \in G \mid f(x) = y\}.$$

If y is not in the image of f, then $f^{-1}(y)$ is empty; meanwhile if y *is* in the image of f, then we've seen that $f^{-1}(y)$ is a coset of $\ker(f)$. Then applying what we've learned about cosets gives the following corollary:

Corollary 1.68

For any homomorphism f, we have $[G : \ker(f)] = |\text{im}(f)|$, or equivalently, $|G| = |\ker(f)| \cdot |\text{im}(f)|$.

§1.9 Normal Subgroups

So far, we've only worked with *left* cosets. But we can define *right* cosets in the exact same way:

Definition 1.69. Given a subgroup $H \leq G$, a *right coset* of H is a subset $Ha = \{ha \mid h \in H\}$.

The same results we've seen for left cosets all apply to right cosets as well — the size of any right coset is $|H|$, and the right cosets of H always partition G.

Example 1.70

In the group S_3, find the *right* cosets of the subgroup $\langle y \rangle$.

Solution. The right cosets are $\{e, y\}$, $\{x, yx\} = \{x, x^2y\}$, and $\{x^2, yx^2\} = \{x^2, xy\}$. □

Note that these cosets still partition S_3, but this is a *different* partition than the one we got from the left cosets $\{e, y\}$, $\{x, xy\}$, and $\{x^2, x^2y\}$.

Claim 1.71 — There is a bijection between left and right cosets, given by taking the inverse — if C is a left coset, then the set $C^{-1} = \{x^{-1} \mid x \in C\}$ is a right coset.

Proof. Let $C = aH$. Then we have $(ah)^{-1} = h^{-1}a^{-1}$ for any h. But h is in H if and only if h^{-1} is, so then C^{-1} is exactly the right coset Ha^{-1}. □

Question 1.72. For which subgroups $H \leq G$ do the left and right cosets give the *same* partition of G?

Subgroups with this property are quite important, so they have a name:

Definition 1.73. A subgroup $H \leq G$ is a *normal subgroup* if $xH = Hx$ for all $x \in G$.

Note that we don't need to consider the case where $xH = Hy$ for *different* elements x and y — then since $x \in Hy$ we would have $Hy = Hx$.

It is sometimes convenient to rewrite the condition as $H = xHx^{-1}$ for all $x \in G$. So a normal subgroup can also be thought of as one which is preserved under *conjugation* by every element $x \in G$.

Proposition 1.74

For any homomorphism f, $\ker(f)$ is normal.

Proof. Let k be an element in $\ker(f)$, so $f(k) = e$. But then for any x, we have

$$f(xkx^{-1}) = f(x)f(k)f(x^{-1}) = f(x)f(x^{-1}) = f(xx^{-1}) = e.$$

So conjugating any element of $\ker(f)$ by x still produces an element of $\ker(f)$. □

In fact, it turns out the converse is true as well:

Fact 1.75 — Every normal subgroup is the kernel of some homomorphism.

We'll prove this in a later class.

Example 1.76

In S_3, the subgroup $\langle x\rangle$ is normal.

Proof. It's possible to check this explicitly, but we don't have to — $\langle x\rangle = \{e, (123), (132)\}$ is the kernel of sgn, so it must be normal. □

Of course, in an abelian group, *every* subgroup is normal.

§1.10 The Correspondence Theorem

Consider a homomorphism $f\colon G \to G'$. As mentioned earlier, we'd like to use homomorphisms to help us understand more complicated groups in terms of simpler ones.

Question 1.77. How are the subgroups of G and G' related?

Given any subgroup $H \le G$, we can take its image $f(H)$, which is a subgroup of G. This lets us go from subgroups of G to those of G'.

Meanwhile, given any subgroup $H' \le G'$, we can take its pre-image

$$f^{-1}(H') = \{x \in G \mid f(x) \in H'\}.$$

This is a subgroup of G — if x and y are both in $f^{-1}(H')$, then $f(x)$ and $f(y)$ are in H'; then $f(xy) = f(x)f(y) \in H'$ as well, so xy is also in $f^{-1}(H')$. So this lets us go from subgroups of G' to subgroups of G.

Question 1.78. Is this correspondence a bijection?

Of course, the answer is no — for example, G could be the trivial group, and G' could be huge.

There are a few constraints we can see immediately — first, $f(H)$ is always contained in $\operatorname{im}(f)$, so the only subgroups of G' we can produce from subgroups of G are the ones in $\operatorname{im}(f)$. Similarly, $f^{-1}(H)$ must always contain $\ker(f)$, so we generally can't produce *all* subgroups of G, only the ones containing $\ker(f)$.

But it turns out that these are essentially the *only* things that can go wrong in the correspondence, and refining the question to account for them gives us the Correspondence Theorem:

Theorem 1.79 (Correspondence Theorem)

Let $f\colon G \to G'$ be a surjective homomorphism. Then there is a bijection between the subgroups of G containing $\ker(f)$ and the subgroups of G'.

Proof. We use the same map from before — to go from subgroups of G to subgroups of G' we take the image, and to go from subgroups of G' to subgroups of G we take the pre-image. Then we want to check that these maps are inverses of each other.

Let $K = \ker(f)$. Then the two directions we need to check are that if we start with a subgroup $K \le H \le G$ then $f^{-1}(f(H)) = H$, and that if we start with $H' \le G'$ then $f(f^{-1}(H')) = H'$. We'll only check the first direction, as the second is similar.

First, $f^{-1}(f(H))$ is the set of all elements in G such that $f(x) \in f(H)$, so clearly $H \subset f^{-1}(f(H))$ by definition.

On the other hand, we have $f(x) = f(h)$ if and only if $x = hk$ for some $k \in K$, meaning that x is in the same coset of K as h is. But since K is contained in H, this means x must be in H as well. So all elements of $f^{-1}(f(H))$ are also elements of H, and therefore we must have $f^{-1}(f(H)) = H$. □

This means if we start off with a complicated group G and we find a surjection from G to a simpler group G', we can use the subgroups of G' to understand the subgroups of G. This idea will come up often, especially in **18.702** in the spring.

Example 1.80

Consider the homomorphism $\mathbb{C}^\times \to \mathbb{C}^\times$ given by $z \mapsto z^2$ (note that this is a homomorphism because $\mathbb{C}^\times$ is abelian, but this isn't generally a homomorphism for a nonabelian group). It's surjective because all complex numbers have a square root, and its kernel is $\{\pm 1\}$.

Then for example, the subgroup $\mathbb{R}^\times$ of the left group corresponds to its image, the subgroup $\mathbb{R}_{>0}$ of the right group.

Meanwhile, the subgroup $\{\pm 1, \pm i\}$ of the right group corresponds to its pre-image, the subgroup $\{e^{2\pi i a/8}\}$ of the left group (consisting of all eighth roots of unity).

§1.11 Quotient Groups

Recall that a subgroup $H \leq G$ is normal if it is preserved by conjugation by any element of G, or in other words, $xHx^{-1} = H$ for all $x \in G$.

Notation 1.81. The notation $H \trianglelefteq G$ is sometimes used to denote that H is a normal subgroup of G.

Earlier, we've seen that for any homomorphism f, its kernel is always normal. We can ask whether the converse is true:

Question 1.82. Given a normal subgroup $N \trianglelefteq G$, does there exist a homomorphism f with $\ker(f) = N$?

We'll see that the answer is yes — we'll construct a new group G' and a homomorphism $G \to G'$ whose kernel is N. First, as a motivating example:

Example 1.83

Consider the normal subgroup $2\mathbb{Z}$ of $\mathbb{Z}$. Then we can take the homomorphism from $\mathbb{Z}$ to $\mathbb{Z}/2\mathbb{Z}$ (the integers mod 2) sending $x \mapsto x \pmod 2$.

So in general, we'd like to construct a version of the integers mod 2.

Note that if $N = \ker(f)$ for some homomorphism f, then each element of $\text{im}(f)$ corresponds to a coset of N — each coset of N is mapped to a different point in $\text{im}(f)$. So we can try to take the image to *be* the cosets of N — let G' be the set of cosets of N in G.

Definition 1.84. Given two cosets C_1 and C_2 of G, their product is $C_1 \cdot C_2 = \{y_1 \cdot y_2 \mid y_1 \in C_1, y_2 \in C_2\}$.

Proposition 1.85

If C_1 and C_2 are cosets of a *normal* subgroup, then so is $C_1 \cdot C_2$.

Proof. Let $C_1 = aN$ and $C_2 = bN$; then we'll show that $C_1 \cdot C_2 = abN$. First, it's clear that $abN \subset C_1 \cdot C_2$, by simply taking $y_1 = a$.

Now for the other direction, take elements an_1 and bn_2 in C_1 and C_2 (with $n_1, n_2 \in N$), so we want to check that $an_1 \cdot bn_2$ is in abN. But since N is normal, we have $bN = Nb$, so $n_1 b$ can be rewritten as bn_3 for some

$n_3 \in N$. Then we have

$$an_1bn_2 = abn_3n_2 \in abN$$

since $n_3n_2 \in N$, as desired. □

Remark 1.86. It's important that N is normal here — as a counterexample when N is not normal, take the group S_3, and the subgroup $H = \{e, y\}$. Then we have $xH = \{x, xy\}$, so

$$xH \cdot xH = \{x^2, x^2y, xyx, xyxy\} = \{x^2, x^2y, y, e\}.$$

This is not a coset of H since it has four elements, not two.

So we can take the product of two cosets, which lets us put a group structure on the set of cosets!

Definition 1.87. Given a normal subgroup $N \trianglelefteq G$, define the *quotient group* G/N as the set of cosets of N, with the group operation $[C_1] \cdot [C_2] = [C_1 \cdot C_2]$.

Theorem 1.88

The quotient G/N is a group, and there exists a surjective homomorphism $\pi\colon G \to G/N$ sending x to the coset containing x, whose kernel is exactly N.

Proof. The fact that G/N is a group is quite straightforward once we know that the operation makes sense. The identity of G/N is $[N]$, since $[aN] \cdot [eN] = [aN]$ for any coset $[aN]$. The inverse of $[aN]$ is $[a^{-1}N]$, since we showed that $[aN] \cdot [a^{-1}N] = [aa^{-1}N] = [N]$. (Note that in *general* the inverse of a left coset is a *right* coset, but here the left and right cosets are the same.) Finally, associativity follows directly from associativity of multiplication in G. So G/N really is a group.

Now π is the map $x \mapsto xN$; the fact that π is a group homomorphism follows directly from the fact that $[xN] \cdot [yN] = [xyN]$. Meanwhile, its kernel is the subset of G which is mapped to $[N]$; this is exactly N. □

Remark 1.89. Note that most of this proof was nearly tautological — most of the work was showing that our multiplication operation really makes sense.

So this answers our question — we've produced a group homomorphism with kernel N. This construction is also useful because if we start with a group and a normal subgroup, we can use it to produce a *new* group:

Example 1.90

The group $\mathrm{SL}_2(\mathbb{R})$ has the normal subgroup $\{\pm I\}$. We can then construct a new group by quotienting out $\mathrm{SL}_2(\mathbb{R})$ by $\{\pm I\}$; this new group is called $\mathrm{PSL}_2(\mathbb{R})$. Note that $\mathrm{PSL}_2(\mathbb{R})$ isn't really a group of matrices — it's a group of matrices up to multiplying by ± 1.

There's another perspective on G/N — given a group G and a subgroup N, we can say that $a \equiv b \pmod{N}$ if a and b lie in the same coset of N. This is an *equivalence relation*, meaning that it satisfies certain axioms. Then when we turn G/N into a group, we're saying that the group operation behaves well under equivalence — if $a \equiv b$ and $c \equiv d$, then $ac \equiv bd$. This lets us put a group structure on the set of *congruence classes.* As a familiar example, this is how modular arithmetic works, using the normal subgroup $n\mathbb{Z} \leq \mathbb{Z}$.

We've now seen how to construct a homomorphism given a normal subgroup. But earlier, we also saw how to construct a normal subgroup given a homomorphism — so now suppose we *started* with a surjective homomorphism $f\colon G \to G'$, which produces the normal subgroup $K = \ker(f)$. Then we can try to feed K into our new construction — we get another surjective homomorphism $\pi : G \to G/K$. But it turns out we essentially haven't done anything, and this new homomorphism is essentially the same as the one we started with. More precisely:

> **Fact 1.91** (First Isomorphism Theorem) — We have $G/\ker(f) \cong G'$.

The isomorphism $\widetilde{f}\colon G/\ker(f) \to G'$ is just given by $\widetilde{f}([xK]) = f(x)$, and in particular $f = \widetilde{f} \circ \pi$. Although this is called a theorem, we should think of it instead as a check that the quotient construction isn't something crazy — there's a correspondence between elements of the image and cosets of the kernel, and that's all that this isomorphism is. The only part of the claim that has content is that this correspondence is compatible with the group structure on both sides.

§2 Linear Maps

We'll now pivot to discussing linear algebra.

§2.1 Vector Spaces

Definition 2.1. A *field* is a set with two operations + and × which satisfy all the rules we'd expect — the operations satisfy associativity and distributivity, all elements form an abelian group under addition, and all nonzero elements form an abelian group under multiplication.

Example 2.2

$\mathbb{C}$, $\mathbb{R}$, and $\mathbb{Q}$ are all fields; $\mathbb{Z}$ is not a field, since we can't generally divide by any nonzero integers — in other words, most nonzero integers don't have multiplicative inverses.

Example 2.3

$\mathbb{Z}/p\mathbb{Z}$ is a field, denoted $\mathbb{F}_p$.

Proof. It's enough to show that every nonzero a has a multiplicative inverse mod p. But we know that $\gcd(a, p) = 1$, so by Bezout's Theorem there exist integers r and s such that $ar + ps = 1$. Then $ar \equiv 1 \pmod{p}$, so r is the inverse of a. □

On the other hand, $\mathbb{Z}/n\mathbb{Z}$ is not a field for composite n — numbers which aren't relatively prime to n don't have inverses.

Definition 2.4. A *vector space* V over a field F is a set with two operations: addition, such that $(V, +)$ is an abelian group, and *scaling*: a map $F \times V \to V$ mapping $(a, \vec{v}) \mapsto a\vec{v}$, satisfying the usual axioms.

Example 2.5

Some examples of vector spaces:

(1) The space of column vectors $(a_1, \ldots, a_n)$ with $a_i \in F$ for all i, denoted F^n, is a vector space.

(2) For any $m \times n$ matrix A with entries in F, the solutions to $A\vec{v} = \vec{0}$ form a vector space.

(3) Linear homogeneous ODEs (ordinary differential equations) form a vector space.

For most things we'll do in linear algebra, it's possible to work over *any* field, instead of just $\mathbb{R}$ — we just need to be able to divide. So for example, we could work with $\mathrm{GL}_n(\mathbb{F}_p)$ instead of $\mathrm{GL}_n(\mathbb{R})$ — this is now a *finite* group.

§2.2 Linear Combinations

Definition 2.6. Given vectors $\vec{v_1}$, …, $\vec{v_n}$ in V, a *linear combination* of these vectors is a vector of the form

$$\vec{v} = a_1\vec{v_1} + \cdots + a_n\vec{v_n}$$

for some $a_1, \ldots, a_n \in F$.

Definition 2.7. For a set $S = \{\vec{v_1}, \dots, \vec{v_n}\}$, the *span* of S, denoted $\mathrm{Span}(S)$, is the set of all vectors $\vec{v}$ which are linear combinations of $\vec{v_1}$, …, $\vec{v_n}$.

Note that $\mathrm{Span}(S)$ is a *vector subspace* of V — it's also a vector space.

We say that a set $S = \{\vec{v_1}, \dots, \vec{v_n}\}$ spans V if $\mathrm{Span}(S) = V$, or in other words, if *every* vector in V can be written as a linear combination of the vectors $\vec{v_i}$.

Definition 2.8. A set of vectors $\vec{v_1}$, …, $\vec{v_n}$ are *linearly independent* if the only $(a_1, \dots, a_n)$ for which

$$a_1\vec{v_1} + \cdots + a_n\vec{v_n} = 0$$

is $a_1 = \cdots = a_n = 0$.

Equivalently, $\vec{v_1}$, …, $\vec{v_n}$ are linearly independent if and only if there is only one way to write each $\vec{v}$ as a linear combination — if there were two ways to write $\vec{v}$, then we could subtract them to get a nontrivial solution to $a_1\vec{v_1} + \cdots + a_n\vec{v_n} = 0$.

Definition 2.9. If a set S of vectors both spans V and is linearly independent, then S is a *basis* of V.

If S is a basis of V, then every vector can be written *uniquely* as a linear combination of its elements — there is a unique way to write

$$\vec{v} = a_1\vec{v_1} + \cdots + a_n\vec{v_n}$$

for any vector $\vec{v}$. Then $(a_1, \dots, a_n)$ are called the *coordinates* of $\vec{v}$ in this basis.

Example 2.10

In the vector space $\mathbb{R}^2$, the set $\{(1,1)^\intercal, (3,2)^\intercal, (2,1)^\intercal\}$ spans $\mathbb{R}^2$, but is not linearly independent. But if we remove the last vector, then the set $\{(1,1)^\intercal, (3,2)^\intercal\}$ still spans $\mathbb{R}^2$ and is now linearly independent; so it forms a basis of $\mathbb{R}^2$.

In linear algebra, a common theme is that a good choice of basis can make the problem easier.

We say V is *finite-dimensional* if there exists a finite list of vectors which spans V. For a finite-dimensional vector space, we'd like to actually *define* its dimension — for example, $\mathbb{R}^2$ should have dimension 2. For this, we need the following lemma:

Lemma 2.11

If we have a set $S = \{\vec{v_1}, \dots, \vec{v_r}\}$ which spans V, and a set $L = \{\vec{w_1}, \dots, \vec{w_s}\}$ which is linearly independent, then:

(1) We can remove elements from S to produce a basis of V.

(2) We can add elements of S to L to produce a basis of V.

(3) We have $|S| \geq |L|$, or in other words $r \geq s$.

Corollary 2.12

If S and L are both bases of V, then they have the same number of vectors.

This lets us define the dimension:

Definition 2.13. The *dimension* of V is the number of vectors in a basis of V.

Proof of Lemma 2.11. First we'll prove (1). If S is linearly independent then we're done; otherwise, we have

$$a_1\vec{v_1} + \cdots + a_r\vec{v_r}$$

for some scalars a_i which are not all zero. Without loss of generality $a_r \neq 0$; then we have

$$\vec{v_r} = -a_r^{-1}(a_1\vec{v_1} + \cdots + a_{r-1}\overrightarrow{v_{r-1}}).$$

This means $\vec{v_r}$ is in the span of the remaining vectors $\vec{v_1}$, ..., $\overrightarrow{v_{r-1}}$; therefore we can remove it from S without changing $\text{Span}(S)$, since any occurrence of $\vec{v_r}$ can be replaced with this expression.

So we've deleted one vector from S, without decreasing its span. We can keep doing this until the vectors are all linearly independent; this must happen at some point, because we can't delete all the vectors.

Now we'll prove (2). First, if S is contained in $\text{Span}(L)$, then since S spans V, so must L; then L is a basis and we're done. So now assume $\vec{v_1}$ is in S but not in $\text{Span}(L)$. Then add $\vec{v_1}$ to L.

Then we claim L is still linearly independent — if we have

$$c\vec{v_1} + a_1\vec{w_1} + \cdots + a_s\vec{w_s} = 0,$$

then we must have $c = 0$ since otherwise $\vec{v_1}$ would have been in $\text{Span}(L)$, and then we must have $a_1 = \cdots = a_s = 0$ since the original set was linearly independent.

We can keep on adding vectors from S to L until we're stuck and L spans V. This must happen at some point — at the least, if we've added everything in S to L then L definitely spans V. Now since L spans V and is still linearly independent, it is a basis for V.

Finally we'll prove (3). Since S spans V, we can write each $\vec{w_j}$ as a linear combination of the $\vec{v_i}$. Let

$$\vec{w_j} = \sum_{i=1}^{r} a_{ij}\vec{v_i}$$

for each $1 \leq j \leq s$. Let A be the $r \times s$ matrix consisting of the a_{ij}; then in matrix notation, this system of equations becomes

$$\begin{bmatrix}\vec{w_1} & \cdots & \vec{w_s}\end{bmatrix} = \begin{bmatrix}\vec{v_1} & \cdots & \vec{v_r}\end{bmatrix} A.$$

Now assume for contradiction that $r < s$. Then the linear system $A\vec{x} = 0$ has more variables than equations, so it must have a nontrivial solution for $\vec{x}$ (for example, this can be proven by putting A into row echelon form). But then we have

$$\sum x_i\vec{w_i} = \begin{bmatrix}\vec{w_1} & \cdots & \vec{w_s}\end{bmatrix}\vec{x} = \begin{bmatrix}\vec{v_1} & \cdots & \vec{v_r}\end{bmatrix} A\vec{x} = 0.$$

This contradicts the fact that the $\vec{w_i}$ are linearly independent; so we must have $r \geq s$. □

§2.3 Linear Transformations

Definition 2.14. Given vector spaces V and W, a *linear transformation* between them is a map $T\colon V \to W$ which is compatible with the vector space operations: we have $T(\vec{v_1} + \vec{v_2}) = T(\vec{v_1}) + T(\vec{v_2})$ for all vectors $\vec{v_1}$ and $\vec{v_2}$, and $T(a\vec{v}) = aT(\vec{v})$ for all vectors $\vec{v}$ and scalars a.

Note that V and W must be vector spaces over the *same* field — when doing linear algebra, we generally fix the field at the very beginning.

Definition 2.15. An *isomorphism* between vector spaces is a bijective linear transformation.

Similarly to the case of group isomorphisms, we can check that if $T\colon V \to W$ is a bijective linear transformation, then its inverse (which exists because it is a bijection) is also a linear transformation.

Example 2.16

Given a set $S = \{\vec{v_1}, \dots, \vec{v_n}\}$ of vectors in V, we can define the transformation $T_S \colon F^n \to V$ as

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \mapsto \sum_{i=1}^{n} a_i \vec{v_i}.$$

In fact, we can rewrite all the properties of a set of vectors described into the previous section as properties of this map — S is linearly independent if and only if T_S is injective, and S spans V if and only if T_S is surjective. In particular, S is a basis for V if and only if T_S is an isomorphism.

In order to describe a linear transformation T, it's enough to describe what T does to a basis of V — then we can use the fact that T interacts well with linear combinations in order to calculate $T(\vec{v})$ for *any* $\vec{v}$.

§2.3.1 Coordinates and Change of Basis

Given a basis $\vec{w_1}$, …, $\vec{w_n}$ of V, we can define a linear transformation $B \colon F^n \to V$ sending $\vec{e_i} \mapsto \vec{w_i}$ for each i. This is an isomorphism (as described in the example), so it has an inverse. In particular, if $B^{-1}(\vec{w}) = (a_1, \dots, a_n)^\intercal$, then we have $\vec{w} = a_1\vec{w_1} + \cdots + a_n\vec{w_n}$, so $B^{-1}(\vec{w})$ gives the *coordinates* of $\vec{w}$ in the basis consisting of $\vec{w_1}$, …, $\vec{w_n}$.

Linear transformations and matrices are closely related, as mentioned much earlier. Given a $m \times n$ matrix A, we can define a linear transformation $T \colon F^n \to F^m$ sending $\vec{v} \mapsto A\vec{v}$. On the other hand, given the transformation T we can uniquely recover the corresponding matrix A — the columns of A should be $T(\vec{e_1})$, …, $T(\vec{e_n})$. This means matrices and linear transformations are essentially the same thing — in fact, matrices and linear transformations both form vector spaces, and this correspondence is an isomorphism of vector spaces. Then a linear transformation $T \colon F^n \to F^m$ is an isomorphism if and only if the corresponding matrix A is invertible, meaning that $m = n$ and $A \in \mathrm{GL}_n(F)$.

Question 2.17. If V is finite-dimensional, how can we relate two bases of V?

Suppose we have one basis $\{\vec{v_1}, \dots, \vec{v_n}\}$, which provides an isomorphism $B \colon F^n \to V$, and another basis $\{\vec{w_1}, \dots, \vec{w_n}\}$, which also provides an isomorphism $B' \colon F^n \to V$. We can then set $P = B^{-1} \circ B'$, which is an isomorphism from $F^n \to F^n$ such that $B' = B \circ P$.

$$\begin{tikzcd}
 & V & \\
F^n \arrow[ur, "B"] & & F^n \arrow[ul, "B'"'] \arrow[ll, "P"]
\end{tikzcd}$$

Since P is an isomorphism $F^n \to F^n$, it corresponds to a $n \times n$ invertible matrix; we'd like to figure out the contents of this matrix. In order to figure out the columns of P, we want to figure out what P does to the standard basis vectors $\vec{e_i}$. We have

$$P(\vec{e_i}) = B^{-1}(B'(\vec{e_i})) = B^{-1}(\vec{w_i}),$$

which is the coordinates of $\vec{w_i}$ in the basis $\{\vec{v_1}, \dots, \vec{v_n}\}$. So the columns of P are exactly the coordinates of the $\vec{w_i}$ in terms of the $\vec{v_j}$. Similarly, the columns of P^{-1} are the coordinates of the $\vec{v_i}$ in terms of the $\vec{w_j}$.

Now suppose we have *any* vector $\vec{v} \in V$, and we write down its coordinates in both bases — let $\vec{x} \in F^n$ be the coordinates of $\vec{v}$ in the basis B, and $\vec{y} \in F^n$ be the coordinates of $\vec{v}$ in the basis B'. Then we have $\vec{x} = B^{-1}(\vec{v})$ and $\vec{y} = B'^{-1}(\vec{v})$, so since $P = B^{-1} \circ B'$, we have $P(\vec{y}) = \vec{x}$, and conversely $P^{-1}(\vec{x}) = \vec{y}$.

By choosing a basis, we can write down everything in terms of coordinates. Suppose we have a linear transformation $T\colon V \to W$. Then we can pick a basis $B = \{\vec{v_1}, \dots, \vec{v_n}\}$ for V, and a basis $C = \{\vec{w_1}, \dots, \vec{w_m}\}$ for W. This gives an isomorphism $A = C^{-1} \circ T \circ B$ from $F^n \to F^m$, which corresponds to a $m \times n$ matrix. So the things we study about matrices in F^n really allow us to study *any* linear transformation (assuming that our vector spaces are finite-dimensional).

Example 2.18

Let V be the set of complex functions satisfying $f''(t) = f(t)$, and W the set of complex functions satisfying $f''(t) = -f(t)$. Define the linear transformation $T\colon V \to W$ as $f(t) \mapsto f(it)$.

One basis for V is $\{e^t, e^{-t}\}$, and one basis for W is $\{\cos t, \sin t\}$. Now to find the corresponding matrix A, its columns are given by the coordinates of $T(\vec{v_i})$ in the basis $\vec{w_j}$. We have $e^t \mapsto e^{it} = \cos t + i \sin t$, so the coordinates of its image are $(1, i)$. Similarly, $e^{-t} \mapsto e^{-it} = \cos t - i \sin t$, so its coordinates are $(1, -i)$. So we have

$$A = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}.$$

We could have produced a different matrix by choosing a different basis — for example, the basis $\{e^{it}, e^{-it}\}$ for W would produce

$$A' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

In the above example, we saw that by changing which basis we used, we could make A into the identity matrix. So we can ask how "nice" we can make the matrix in general:

Question 2.19. Can we choose bases B and C so that the matrix A looks very nice?

We'll come back to this later.

§2.3.2 The Dimension Formula

Definition 2.20. Given a linear transformation $T\colon V \to W$, its *kernel* is the set of vectors $\vec{v}$ in V for which $T(\vec{v}) = 0$, and its *image* is the set of vectors $\vec{w}$ in W such that $\vec{w} = T(\vec{v})$ for some $\vec{v}$.

The kernel and image are vector subspaces of V and W respectively, by the same reasoning as in the case of groups. So it makes sense to define their dimensions:

Definition 2.21. The dimension of $\operatorname{im}(T)$ is called the *rank*, and the dimension of $\ker(T)$ is called the *nullity*.

Equivalently, given a matrix A, its rank is the dimension of the span of its columns.

Theorem 2.22 (Dimension Formula)

We have $\dim \ker(T) + \dim \operatorname{im}(T) = \dim V$.

This is also known as the Rank-Nullity Theorem.

Remark 2.23. This is reminiscent of the formula $|G| = |\ker(f)| \cdot |\mathrm{im}(f)|$ from group homomorphisms.

Proof. Pick a basis $\vec{v_1}$, …, $\vec{v_k}$ of $\ker(T)$. Since this set of vectors is linearly independent, we can extend it to a basis of V, by adding some vectors $\overrightarrow{v_{k+1}}$, …, $\vec{v_n}$.

We know that $T(\vec{v_i}) = 0$ for all $1 \le i \le k$. Meanwhile, let $T(\vec{v_i}) = \vec{w_i}$ for each $k+1 \le i \le n$; then the vectors $\vec{w_i}$ are all in $\mathrm{im}(T)$.

Claim — The vectors $\overrightarrow{w_{k+1}}$, …, $\vec{w_n}$ form a basis for $\mathrm{im}(T)$.

Proof. First, to show they span $\mathrm{im}(T)$, we have

$$\mathrm{im}(T) = \mathrm{Span}(T(\vec{v_1}), \dots, T(\vec{v_n}))$$

by linearity. But $T(\vec{v_1})$, …, $T(\vec{v_k})$ are all 0, so we can remove them without affecting the span.

Now to show they're linearly independent, suppose

$$a_{k+1}\overrightarrow{w_{k+1}} + \cdots + a_n\vec{w_n} = 0.$$

Then using linearity, we can rewrite this as

$$T(a_{k+1}\overrightarrow{v_{k+1}} + \cdots + a_n\vec{v_n}) = 0.$$

But this means $a_{k+1}\overrightarrow{v_{k+1}} + \cdots + a_n\vec{v_n}$ must be in $\ker(T)$, and therefore we can write

$$a_{k+1}\overrightarrow{v_{k+1}} + \cdots + a_n\vec{v_n} = a_1\vec{v_1} + \cdots + a_k\vec{v_k}$$

for some a_1, …, a_k. But since $\vec{v_1}$, …, $\vec{v_n}$ form a basis for V, they must be linearly independent, and therefore all coefficients are 0. □

Now we're done, since $\dim\ker(T) = k$ and $\dim\mathrm{im}(T) = n-k$, and these sum to $\dim V = n$. □

In fact, this proof shows something more. Take a basis $\overrightarrow{v_{k+1}}$, …, $\vec{v_n}$, $\vec{v_1}$, …, $\vec{v_k}$ of V as described. Then take the basis $\overrightarrow{w_{k+1}}$, …, $\vec{w_n}$ of $\mathrm{im}(T)$ as described, and add vectors $\vec{u_1}$, …, $\vec{u_k}$ to extend this to a basis of W. In these bases, it's easy to describe T — it sends $\vec{v_i} \mapsto \vec{w_i}$ for each $k+1 \le i \le n$, and $\vec{v_i} \mapsto 0$ for all other i. So then in the matrix corresponding to T, the first $n-k$ entries on the diagonal are all 1, and all other entries are 0: so T looks like

$$\left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & 1 & \\ \hline \multicolumn{4}{c|}{0} & 0 \end{array}\right].$$

This answers our question from earlier — for *any* linear transformation, we can choose bases to make the transformation have this form, which is very nice. As a special case, suppose we started with a linear transformation $F^n \to F^m$, which *already* corresponds to a matrix M. Then we can choose *new* bases for F^n and F^m in which M has this form; this corresponds to choosing invertible matrices P and Q, and writing the new matrix $A = Q^{-1}MP$.

Corollary 2.24

Given any $m \times n$ matrix M, there exist $P \in \mathrm{GL}_n(F)$ and $Q \in \mathrm{GL}_m(F)$ such that $Q^{-1}MP$ has some number of 1's at the beginning of its diagonal, and 0's everywhere else.

This comes back to a remark made earlier, that the choice of basis can often make your life a lot easier.

Remark 2.25. To explicitly tie this back to the Dimension Formula, columns with all zeros correspond to the kernel, while the columns with a 1 correspond to the image.

Corollary 2.26

Given any $m \times n$ matrix M, the rank of M is the same as the rank of its transpose $M^\intercal$.

In other words, the *column rank* (the dimension of the span of its columns) is the same as the *row rank* (the dimension of the span of its rows).

Proof. Write $A = Q^{-1}MP$ in the form described. Then this is clearly true for A — the rank of A is just the number of 1's on the diagonal, and its transpose has the same number of 1's. But A and M are isomorphic (since we obtained A from M by multiplying by invertible matrices, or equivalently by changing basis), and $A^\intercal$ and $M^\intercal$ are isomorphic. So the rank of A is the same as the rank of M, and the rank of A is the same as the rank of $M^\intercal$; therefore M and $M^\intercal$ have the same rank as well. □

§2.4 Linear Operators

So far, we've looked at linear maps between different spaces. But we can also look at linear maps on a *fixed* space:

Definition 2.27. A *linear operator* is a linear transformation $T\colon V \to V$.

Example 2.28

Some examples of linear operators:

(1) In the vector space $\mathbb{R}^2$, rotation by θ counterclockwise is a linear operator (for any angle θ).

(2) In the vector space of polynomials of degree at most 2, the derivative is a linear operator.

In order to understand a linear operator T, we can still choose a basis for V and write down the matrix corresponding to T. The only difference between this case and that of general linear maps is that now since we have V on both sides, we only need *one* basis instead of two. Then if we have a basis giving an isomorphism $B\colon F^n \to V$, this turns T into a $n \times n$ square matrix A, whose columns are the images of each basis vector under T.

Example 2.29

In our two above examples:

(1) If we take the standard basis for $\mathbb{R}^2$, the corresponding matrix is

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

(2) If we take the basis $\{1, t, t^2\}$, then the derivative maps $1 \mapsto 0$, $t \mapsto 1$, and $t^2 \mapsto 2t$, giving the matrix

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

Proposition 2.30

If V is finite-dimensional, then a linear operator $T\colon V \to V$ is injective if and only if it is surjective.

We'll always assume vector spaces are finite-dimensional unless otherwise stated.

Proof. We can use the dimension formula — we know $\dim\ker(T) + \dim\operatorname{im}(T) = \dim V$. So $\dim\ker(T) = 0$ if and only if $\dim\operatorname{im}(T) = \dim V$, which occurs if and only if $\operatorname{im}(T) = V$. □

So to check that a linear operator is an isomorphism, it's enough to check only *one* of injectivity and surjectivity. In this sense, finite-dimensional vector spaces behave a lot like finite sets.

§2.4.1 Change of Basis

Suppose we have a linear operator $T\colon V \to V$, and a basis $B\colon F^n \to V$ in which T becomes the matrix A. Now suppose we want to write T in a different basis — take an invertible matrix $P\colon F^n \to F^n$ and use the new basis $B' = BP$, and let A' be the matrix of T in this basis.

$$\begin{tikzcd}
V \arrow[r, "T"] & V \\
F^n \arrow[u, "B"'] \arrow[r, "A"] & F^n \arrow[u, "B"] \\
F^n \arrow[u, "P"'] \arrow[uu, bend left, "B'"] \arrow[r, "A'"'] & F^n \arrow[u, "P"] \arrow[uu, bend right, "B'"']
\end{tikzcd}$$

To follow the arrow corresponding to A', we'd first go up using P, then right using A, then down using P^{-1}. So we have $A' = P^{-1}AP$, and therefore changing basis *conjugates* the matrix.

Definition 2.31. Two matrices A and A' are *similar* if there exists an invertible matrix P such that $A' = P^{-1}AP$.

So two matrices are similar if they correspond to the same linear operator written in different bases.

Note that this means, for example, that we can define the determinant of a *linear operator*, rather than just a matrix — given a linear operator $T\colon V \to V$, we can pick a basis of V to produce a square matrix A, and define the determinant of T as the determinant of A. This is well-defined because if we chose a different basis, then we'd get a matrix $A' = P^{-1}AP$ for some invertible P, and then

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

So even though the *matrix* depends on the choice of basis, its determinant doesn't.

Remark 2.32. This suggests that the determinant is intrinsic to T, in some sense. In fact, over $\mathbb{R}$, the determinant has a meaning related to volume. Something similar is true even over other fields (where volume may not make sense).

§2.5 Diagonalization

Question 2.33. How nice can we make the matrix of a linear operator by changing the basis?

Equivalently, given a matrix A, we'd like to find a matrix *similar* to A which is nice. In the case of linear transformations, we could make the matrix *really* nice by changing the basis on both sides. However, here we have less flexibility, since we only get to choose *one* basis.

Example 2.34

Take the matrix

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix},$$

which is a linear operator on $\mathbb{R}^2$. By changing the basis for $\mathbb{R}^2$, how nice can we make the new matrix?

Solution. Note that

$$A\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

and similarly we have

$$A\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -\begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

So the linear operator does something really nice to $(1,1)^\intercal$ and $(-1,1)^\intercal$ — it just stretches (or flips) them. Then if we take $(1,1)^\intercal$ and $(-1,1)^\intercal$ as our basis — meaning we take the change of basis matrix

$$P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},$$

then we get the new matrix

$$A' = P^{-1}AP = \begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix}.$$

So we've produced a *diagonal* matrix! □

§2.5.1 Eigenvectors

Definition 2.35. A nonzero vector $\vec{v}$ is an *eigenvector* for a linear operator T if $T\vec{v} = \lambda\vec{v}$ for some scalar λ, which we call the *eigenvalue* of $\vec{v}$.

So the linear operator T doesn't change the direction of $\vec{v}$ — it just scales $\vec{v}$. So in some sense, eigenvectors are the directions in which T behaves nicely.

The reason we could make A a diagonal matrix in the above example is that we had enough eigenvectors to form a basis.

Definition 2.36. A basis $\vec{v_1}, \dots, \vec{v_n}$ such that each $\vec{v_i}$ is an eigenvector of T is called a *eigenbasis* for T.

If we have an eigenbasis where $T\vec{v_i} = \lambda_i\vec{v_i}$ for each i, then in this basis T becomes the diagonal matrix

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Diagonal matrices are really nice — for example, it's easy to take the nth power of a diagonal matrix.

Definition 2.37. We say T is *diagonalizable* if there exists a basis in which T is diagonal.

Note that a basis in which T is diagonal is exactly an eigenbasis. Of course, we can also discuss whether a square *matrix* is diagonalizable, by the same definition — meaning that it's similar to a diagonal matrix.

Question 2.38. How can we find eigenvectors, eigenvalues, and an eigenbasis?

We can start by trying to find the *eigenvalues*. A scalar λ is an eigenvalue if and only if there exists some nonzero $\vec{v}$ for which $A\vec{v} = \lambda\vec{v}$, or equivalently

$$(A - \lambda I)\vec{v} = 0.$$

So λ is an eigenvalue if and only if $\ker(A - \lambda I)$ is nonzero, meaning that $A - \lambda I$ is not invertible; this occurs exactly when

$$\det(\lambda I - A) = 0.$$

But we can imagine expanding out the determinant; then this is a polynomial equation in λ!

Definition 2.39. The *characteristic polynomial* of A is the degree n polynomial $p_A(t) = \det(tI - A)$.

Proposition 2.40

The characteristic polynomial of a linear operator does not depend on the choice of basis.

In other words, if A and A' are similar, then $p_{A'}(t) = p_A(t)$.

Proof. This follows directly from the fact that the characteristic polynomial is a determinant. More precisely, if $A' = P^{-1}AP$, then $(tI - A') = P^{-1}(tI - A)P$ as well, so their determinants are the same. □

In fact, this observation has another useful corollary — we know that *all* terms of the characteristic polynomial are independent of the choice of basis. But we can write out a few of the terms — if A consists of entries a_{ij}, then

$$p_A(t) = t^n - (a_{11} + \cdots + a_{nn})t^{n-1} + \cdots + (-1)\det A.$$

The quantity $a_{11} + \cdots + a_{nn}$ is called the *trace* of the matrix; this implies that the trace is independent of the choice of basis, or equivalently, that $\operatorname{tr}(P^{-1}AP) = \operatorname{tr}(A)$ for any matrix A and invertible matrix P.

Returning to the problem at hand, we know the eigenvalues are exactly the roots of the characteristic polynomial. Unfortunately, it's possible that there *are* no roots:

Example 2.41

Consider the linear operator on $\mathbb{R}^2$ given by rotation by θ. This has characteristic polynomial

$$p_A(t) = \det \begin{bmatrix} t - \cos\theta & \sin\theta \\ -\sin\theta & t - \cos\theta \end{bmatrix} = t^2 - 2\cos\theta + 1,$$

which has no real roots unless θ is a multiple of π. So this operator has *no* eigenvalues or eigenvectors. (This is unsurprising because when we rotate every vector, we don't preserve any vector's direction.)

This is a problem, but we can fix it by working over $\mathbb{C}$ instead of $\mathbb{R}$. Then every polynomial factors as a product of linear terms, so a polynomial of degree n has n roots (with multiplicity). For the rest of this discussion, we'll assume we're working over the field $\mathbb{C}$ to take care of this first obstacle.

Question 2.42. Can we always find an eigenbasis?

Unfortunately, the answer is no — we may not be able to find enough linearly independent eigenvectors to form a basis.

Example 2.43

Take the matrix

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Its characteristic polynomial is $p_A(t) = t^2$, so its only eigenvalue is 0 (with multiplicity 2). Then if A were similar to a diagonal matrix, that diagonal matrix would have to be the zero matrix, making A itself the zero matrix; this is a contradiction. So A cannot be diagonalizable.

In the above example, we only had one eigenvalue 0; but this eigenvalue corresponded to a 1-dimensional kernel, not a 2-dimensional kernel. So we couldn't find enough eigenvectors to form a basis.

Remark 2.44. We'll see later that this is the most important counterexample, in some sense.

Since repeated roots can potentially cause problems, let's first work with eigenvectors of *different* eigenvalues.

Proposition 2.45

Suppose a $n \times n$ matrix has k eigenvectors $\vec{v_1}, \ldots, \vec{v_k}$ with corresponding eigenvalues $\lambda_1, \ldots, \lambda_k$ which are all distinct. Then $\vec{v_1}, \ldots, \vec{v_k}$ are linearly independent.

Proof. We'll use induction on k. The base case $k = 1$ is clearly true, since the only way for a set of one vector to not be linearly independent is if it is the zero vector, and eigenvectors are nonzero by definition.

Now assume it's true for $k-1$ eigenvectors, and we'll show it's true for k. Suppose we have a linear relation

$$a_1\vec{v_1} + \cdots + a_k\vec{v_k} = 0.$$

Now apply A to both sides. Since multiplication by A is linear, we get

$$a_1 \cdot A\vec{v_1} + \cdots + a_k \cdot A\vec{v_k} = 0,$$

and since each $\vec{v_i}$ is an eigenvector, this means

$$a_1\lambda_1\vec{v_1} + \cdots + a_k\lambda_k\vec{v_k} = 0.$$

So we started off with one linear combination that resulted in 0, and now we've produced another one! We can now scale the original linear combination by λ_k and subtract to get

$$a_1(\lambda_1 - \lambda_k)\vec{v_1} + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k)\overrightarrow{v_{k-1}} = 0.$$

But this is a linear relation between $\vec{v_1}, \ldots, \overrightarrow{v_{k-1}}$, so by the inductive hypothesis, all the coefficients must be 0. But we can't have $\lambda_i - \lambda_k = 0$ for any k, so this means $a_1, \ldots, a_{k-1}$ are all 0. But then $a_k\vec{v_k} = 0$, and since $\vec{v_k}$ is nonzero, this means a_k is 0 as well. □

Corollary 2.46

If the characteristic polynomial of A factors as

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

where all the λ_i are distinct (in other words, it has no repeated roots), then A has an eigenbasis and is diagonalizable.

Proof. Let $\vec{v_1}, \ldots, \vec{v_n}$ be eigenvectors corresponding to $\lambda_1, \ldots, \lambda_n$. Then by the above proposition, $\vec{v_1}, \ldots, \vec{v_n}$ must be linearly independent; then their span has dimension n, so they must form a basis. □

This means if the characteristic polynomial has no repeated roots, then we immediately know the matrix is diagonalizable, without even having to compute the eigenvectors. This is quite strong — *most* of the time, the characteristic polynomial will not have repeated roots (more precisely, the set of matrices for which there are repeated roots has measuzre 0).

In general, suppose the characteristic polynomial factors as

$$p_A(t) = (t - \lambda_1)^{e_1}\cdots(t - \lambda_k)^{e_k},$$

where the λ_i are all distinct. Then we can find the vector spaces $V_{\lambda_i} = \ker(\lambda_i I - A)$. Each of these spaces has dimension at least 1 (every eigenvalue has at least one eigenvector), so we can produce a basis for each one. (This is computationally easy — given an explicit matrix, we can find a basis for its kernel by using row operations.)

Then using Proposition 2.45, our set of all these basis vectors is linearly independent as well. So if we can find n basis vectors in total, then they form an eigenbasis and A is diagonalizable; but if we can't, then A is not diagonalizable. We'll see later that $\dim V_{\lambda_i} \leq e_i$ for each i, so diagonalization fails if one of these bounds is strict.

§2.5.2 Jordan Normal Form

We've seen that *almost* all matrices are diagonalizable, but we'd still like to figure out the nicest form we can put a matrix into even if it's not diagonalizable.

Definition 2.47. Given a positive integer a and a scalar λ, the *Jordan block* $J_a(\lambda)$ is the $a \times a$ matrix with λ's on the diagonal, 1's directly above the diagonal, and 0's everywhere else: so we have

$$J_a(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}.$$

This matrix is not diagonalizable — its characteristic polynomial is $(t-\lambda)^a$, so λ is its only eigenvalue, and $\vec{e_1}$ is the only λ-eigenvector. So if $a > 1$, then $J_a(\lambda)$ doesn't have an eigenbasis.

So the matrices $J_a(\lambda)$ can't be diagonalized, but in some sense they capture everything that can go wrong when attempting to diagonalize (assuming we're working over the field $\mathbb{C}$):

Theorem 2.48

Given a linear operator $T\colon V \to V$ with $\dim V = n$, there exists a basis for V and some (a_1, λ_1), ..., (a_r, λ_r) such that the matrix corresponding to T is the block-diagonal matrix formed by concatenating $J_{a_1}(\lambda_1)$, ..., $J_{a_r}(\lambda_r)$ along the diagonal.

So this answers our question about how nice we can make a linear operator — we can find a basis in which it looks like a bunch of Jordan blocks glued together. In fact, these Jordan blocks are unique up to rearrangement, and this decomposition is called the *Jordan decomposition* of T.

Example 2.49

What are the possible Jordan decompositions when $n = 4$?

Solution. We must have $a_1 + \cdots + a_r = 4$. The ways to have positive integers summing to 4 are 4, $3+1$, $2+2$, $2+1+1$, and $1+1+1$.

If $a_1 = 4$, then the Jordan form of our matrix is

$$\begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}.$$

If $(a_1, a_2) = (3, 1)$, then the Jordan form is

$$\begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{bmatrix}.$$

If $(a_1, a_2) = (2, 2)$, then the Jordan form is

$$\begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{bmatrix}.$$

If $(a_1, a_2, a_3) = (2, 1, 1)$, then the Jordan form is

$$\begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{bmatrix}.$$

Finally, if $(a_1, a_2, a_3, a_4) = (1, 1, 1, 1)$, then the Jordan form is the diagonal matrix

$$\begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix}.$$

□

In particular, a matrix A is diagonalizable if and only if all the Jordan blocks have size 1.

Note that the characteristic polynomial of a matrix in Jordan normal form is

$$p_A(t) = (t - \lambda_1)^{a_1} \cdots (t - \lambda_r)^{a_r}.$$

This is closely related to the factorization $(t-\lambda_1)^{e_1} \cdots (t-\lambda_k)^{e_k}$ we defined earlier, but it's not the same — there may be multiple Jordan blocks corresponding to the same eigenvalue. So it's not generally possible to figure out the Jordan decomposition just from the characteristic polynomial. But we do get *some* information — for each eigenvalue λ, the sizes of the Jordan blocks corresponding to λ sum to the exponent of $t - \lambda$ in the characteristic polynomial.

Every matrix has a Jordan form, but *almost* every matrix is diagonalizable — if we take any matrix and perturb its entries a bit, it will be diagonalizable. So Jordan form is necessary 0 percent of the time; but it's useful to have a result that works for *all* matrices.

To set up the proof of Jordan normal form, let's think about what the Jordan blocks really represent.

Example 2.50

As an operator, the matrix $J_4(0)$ affects the basis vectors by sending

$$\vec{e_4} \mapsto \vec{e_3} \mapsto \vec{e_2} \mapsto \vec{e_1} \mapsto 0.$$

In particular, if we denote this linear operator by T, then $T^4\vec{e_i} = 0$ for all basis vectors $\vec{e_i}$, which means $T^4\vec{x} = 0$ for *all* vectors $\vec{x}$.

In general, if T has a Jordan block with eigenvalue 0, then there is some n such that $T^n\vec{x} = 0$ for all $\vec{x}$ in the corresponding subspace. Similarly, if T has a Jordan block with eigenvalue λ, then there is some n such that $(T - \lambda I)^n\vec{x} = 0$ for all $\vec{x}$ in the corresponding subspace.

In the above example, we had a chain of vectors which eventually reached 0. If for example we had two copies of $J_2(0)$ instead, then we'd have *two* chains — $\vec{e_2} \mapsto \vec{e_1} \mapsto 0$ and $\vec{e_4} \mapsto \vec{e_3} \mapsto 0$.

Definition 2.51. A vector $\vec{x}$ is called a *generalized eigenvector* of T if $(T - \lambda I)^n\vec{x} = 0$ for some n.

So Jordan normal form corresponds to chains of generalized eigenvectors, in some sense. In order to prove Jordan normal form, there's a few more concepts we'll make use of:

Definition 2.52. Given a linear operator $T\colon V \to V$ and a subspace $W \subset V$, we say that W is *T-invariant* if for each $\vec{w} \in W$, we also have $T(\vec{w}) \in W$.

In other words, W is T-invariant if $T(W) \subset W$.

Definition 2.53. Given a vector space V and two subspaces W and W', we say that V is the *direct sum* of W and W' if every vector in V can be written uniquely as the sum of an element of W and one of W'.

In other words, for V to be a direct sum of W and W', we should be able to take a basis of W and one of W', and string them together to get a basis of V. Writing V as a direct sum of W and W' is sometimes also called a *splitting* of V.

Fact 2.54 — If $\dim W + \dim W' = \dim V$ and the only vector W and W' have in common is 0, then V is the direct sum of W and W'.

Proof. The elements of the bases of W and W' must be linearly independent (or else we'd have some nonzero vector in both of them), so they must form a basis for V (since there's the right number of vectors). □

Definition 2.55. A splitting $V = W \oplus W'$ is called T-invariant if both W and W' are T-invariant.

The point of these definitions is that this is essentially what it means for a matrix to be block-diagonal — if a matrix is block-diagonal, then the vector space corresponding to each block is invariant under the matrix, and the direct sum of these vector spaces is V.

Finally, we'll use one more definition:

Definition 2.56. An operator T is *nilpotent* if $T^m = 0$ for some positive integer m.

Note that Jordan blocks $J_a(0)$ are nilpotent (and no other Jordan blocks are).

Now we are ready to prove the existence of a Jordan decomposition.

Proof of Theorem 2.48. We'll induct on $\dim V$. The main idea is to split the vector space into two T-invariant pieces and find a Jordan decomposition for each.

First, we may assume that 0 is an eigenvalue of T — otherwise let λ be some eigenvalue (which must exist), and replace T with $T - \lambda I$. (Then if we get a Jordan decomposition for the new operator $T - \lambda I$, we can get a decomposition for T by simply adding λI.)

Claim — There exists a T-invariant splitting $V = W \oplus U$ such that the operator $T|_W$ is nilpotent, and the operator $T|_U$ is invertible.

The notation $T|_W$ means T restricted to W — so we view T as a linear operator $W \to W$ (which makes sense since W is T-invariant).

Proof. Consider the chain

$$V \supset TV \supset T^2V \supset T^3V \supset \cdots.$$

This gives a nesting family of subspaces. But the dimensions of these spaces form a nondecreasing sequence of nonnegative integers, so they must eventually stabilize; that means eventually the subspaces stop shrinking, and we have

$$T^mV = T^{m+1}V = T^{m+2}V = \cdots.$$

Now define $U = \operatorname{im}(T^m)$ (or in other words, $U = T^mV$) and $W = \ker(T^m)$. We claim that this provides a T-invariant splitting with the claimed properties.

First we'll check that both spaces are invariant. It's clear that U is T-invariant, since $TU = U$; meanwhile W is T-invariant as well since if $\vec{w}$ is in $\ker(T^m)$, then $T^{m+1}\vec{w}$ is zero as well, so $T\vec{w}$ is also in $\ker(T^m)$.

Now $T|_W$ is nilpotent because $(T|_W)^m = 0$ by definition (we defined $W = \ker(T^m)$, so applying T to W for m times will send every vector in W to 0). Meanwhile, $T|_U$ is invertible because $\operatorname{im}(T|_U) = U$, and therefore $T|_U$ must be a bijection.

Finally, it remains to show that $W \oplus U = V$. For this, we'll use Fact 2.54: first, T^m is invertible on U, so no nonzero vector in U can be in $W = \ker(T^m)$. Meanwhile we have

$$\dim W + \dim U = \dim \ker(T^m) + \dim \operatorname{im}(T^m) = \dim V$$

by the Dimension Formula. So by Fact 2.54, we have $W \oplus U = V$. ■

Now we've split V into a nilpotent part and an invertible part. Note that $\dim U < \dim V$, since we assumed that 0 was an eigenvalue of T. So by the inductive hypothesis we can find a Jordan decomposition for $T|_U$, and it's now enough to find one for $T|_W$. So we've reduced the problem to one about *nilpotent* operators.

> **Claim** — If $T\colon V \to V$ is a nilpotent operator, then there is a basis of V in which T acts by chains — meaning that T sends $\vec{e_k} \mapsto \overrightarrow{e_{k-1}} \mapsto \cdots \mapsto \vec{e_1} \mapsto 0$. (There may be several chains.)

Proof. We again use induction on $\dim V$. Set $W = \operatorname{im}(T)$; then W has strictly lower dimension than V (since if $W = V$, then T would be invertible). Then by the inductive hypothesis, we can find a basis for W in which T acts by a bunch of chains:

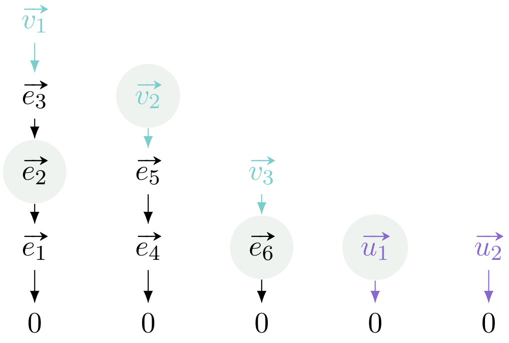
$$\begin{tikzcd}
\vec{e_3} \arrow[d] & & \\
\vec{e_2} \arrow[d] & \vec{e_5} \arrow[d] & \\
\vec{e_1} \arrow[d] & \vec{e_4} \arrow[d] & \vec{e_6} \arrow[d] \\
0 & 0 & 0
\end{tikzcd}$$

Now for each chain, insert the pre-image of the top vector at the beginning of the chain — if there are multiple pre-images of one vector, then choose arbitrarily. Call these vectors $\vec{v_1}$, $\vec{v_2}$, $\dots$, $\vec{v_k}$ — so in this situation we'd have $\vec{v_1} \mapsto \vec{e_3}$, $\vec{v_2} \mapsto \vec{e_5}$, and $\vec{v_3} \mapsto \vec{e_6}$. Additionally, extend the bottom vectors to form a basis of $\ker(T)$, by adding vectors $\vec{u_1}$, $\vec{u_2}$, $\dots$, $\vec{u_\ell}$ which all map to 0.

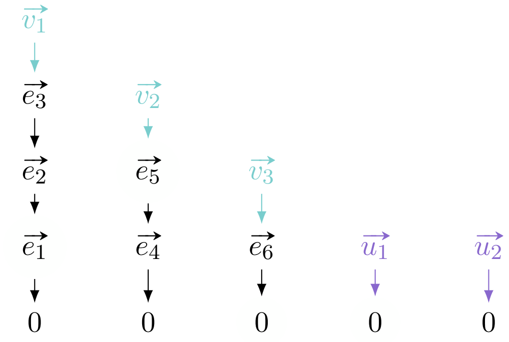
$$\begin{tikzcd}
\vec{v_1} \arrow[d] & & & & \\
\vec{e_3} \arrow[d] & \vec{v_2} \arrow[d] & & & \\
\vec{e_2} \arrow[d] & \vec{e_5} \arrow[d] & \vec{v_3} \arrow[d] & & \\
\vec{e_1} \arrow[d] & \vec{e_4} \arrow[d] & \vec{e_6} \arrow[d] & \vec{u_1} \arrow[d] & \vec{u_2} \arrow[d] \\
0 & 0 & 0 & 0 & 0
\end{tikzcd}$$

We claim that our basis vectors for W, together with these new vectors, form a basis for V. First, there's the right number of them — we started out with $\dim W = \dim \operatorname{im}(T)$ vectors and added in $\dim \ker(T)$ vectors (since the vectors at the bottom of each chain in the new picture form a basis for $\ker(T)$, and we've added one vector in each chain), and we have $\dim V = \dim \operatorname{im}(T) + \dim \ker(T)$ by the Dimension Formula. So it's enough to show that they're linearly independent.

Suppose we have some linear combination of these vectors which equals 0.



Now apply T to this linear combination. This pushes all our vectors down one step in the chain.



Then the terms $a\vec{x}$ for $\vec{x}$ not in $\ker(T)$ are all pushed down one step, meaning that $\vec{x}$ is replaced with a basis vector of W; meanwhile the terms where $\vec{x}$ *is* in $\ker(T)$ all become 0. Then since the basis vectors for W are all linearly independent, the coefficients of all $\vec{x}$ in the first case must be 0. But then our linear combination of the vectors in the second case *also* equals zero, and these vectors are linearly independent as well (since they form a basis for the kernel), which means their coefficients are also all zero. ■

Now returning to our original linear operator, we can find a basis of W in which $T|_W$ acts by chains; then each chain $\vec{e_k} \mapsto \overrightarrow{e_{k-1}} \mapsto \cdots \mapsto \vec{e_1} \mapsto 0$ corresponds to a Jordan block $J_k(0)$. Meanwhile, we can find a basis of U in which $T|_U$ consists of Jordan blocks by the inductive hypothesis; concatenating these gives a Jordan decomposition of T. □

§3 Symmetry

§3.1 Orthogonal Matrices

We'll work over the field $\mathbb{R}$.

Definition 3.1. The *dot product* of two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ is the real number

$$\vec{x} \cdot \vec{y} = \vec{x}^\intercal \vec{y} = \sum_{i=1}^{n} x_i y_i.$$

Fact 3.2 — We have $\vec{x} \cdot \vec{y} = |\vec{x}||\vec{y}|\cos\theta$, where θ is the angle between $\vec{x}$ and $\vec{y}$. In particular, $\vec{x} \cdot \vec{y} = 0$ if and only if $\vec{x}$ and $\vec{y}$ are perpendicular.

Definition 3.3. A basis $\vec{v_1}, \dots, \vec{v_n}$ is *orthonormal* if $\vec{v_i} \cdot \vec{v_j}$ is 1 whenever $i = j$, and 0 whenever $i \neq j$.

In other words, $|\vec{v_i}| = 1$ for all i, and $\vec{v_i} \cdot \vec{v_j} = 0$ for all $i \neq j$.

Definition 3.4. A matrix $A \in \mathrm{GL}_n(\mathbb{R})$ is *orthogonal* if $A\vec{v} \cdot A\vec{w} = \vec{v} \cdot \vec{w}$ for all $\vec{v}$ and $\vec{w}$.

In other words, orthogonal matrices are matrices which preserve the dot product.

There are a few equivalent ways to describe orthogonality:

Theorem 3.5

Given $A \in \mathrm{GL}_n(\mathbb{R})$, the following conditions are all equivalent:

(1) $A\vec{v} \cdot A\vec{w}$ for all $\vec{v}$ and $\vec{w}$.

(2) $|A\vec{v}| = |\vec{v}|$ for all $\vec{v}$.

(3) $A^\intercal A = I$.

(4) The columns of A are an orthonormal basis of $\mathbb{R}^n$.

Note that if A is orthogonal, by (3) its transpose is orthogonal as well; so the *rows* of A are also an orthonormal basis.

Proof. First, (1) implies (2) by taking $\vec{v} = \vec{w}$. On the other hand, we can write

$$\vec{v} \cdot \vec{w} = \frac{1}{2}\left(|\vec{v} + \vec{w}|^2 - |\vec{v}|^2 - |\vec{w}|^2\right),$$

so (2) implies (1) as well — the right-hand side is preserved when we apply A, so $\vec{v} \cdot \vec{w}$ must be preserved by A as well.

To prove (1) and (3) are equivalent, we can write $A\vec{v} \cdot A\vec{w} = \vec{v}^\intercal A^\intercal A\vec{w}$, so then (1) is equivalent to

$$\vec{v}^\intercal A^\intercal A\vec{w}$$

for all $\vec{v}$ and $\vec{w}$. We claim that this is true if and only if $A^\intercal A = I$. It's clear that this is true if $A^\intercal A = I$. On the other hand, if we take $\vec{v} = \vec{e_i}$ and $\vec{w} = \vec{e_j}$, then $\vec{e_i} M \vec{e_j} = m_{ij}$ for any matrix $M = (m_{ij})$, which means all entries of $A^\intercal A$ and I must be the same.

Finally, we'll show that (3) and (4) are equivalent. The condition $A^\intercal A = I$ means that the dot product of the ith row of $A^\intercal$ and the jth row of A is 0 when $i \neq j$ and 1 when $i = j$. But the ith row of $A^\intercal$ is exactly the ith column of A, so this is equivalent to stating that the columns of A form an orthonormal basis. □

Orthogonal matrices can be interpreted geometrically using conditions (1) and (2) — they preserve length, and they preserve angles up to sign (since they must preserve $|\vec{v}|\,|\vec{w}|\cos\theta$, where θ is the angle between any two vectors, and since they preserve length they must then preserve $\cos\theta$ as well).

Notation 3.6. The set of orthogonal matrices is denoted O_n.

Note that O_n is actually a *subgroup* of $\mathrm{GL}_n(\mathbb{R})$ — the conditions of a subgroup can be checked directly using (3). This group is called the *orthogonal group*.

Given a matrix $A \in \mathrm{O}_n$, we can consider its determinant: we have

$$1 = \det(A^\intercal A) = \det(A^\intercal)\det(A) = \det(A^2),$$

so we must have $\det(A) = \pm 1$. So det gives a homomorphism $\mathrm{O}_n \to \{\pm 1\}$. This homomorphism is surjective — for example, we have

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \mapsto 1 \text{ and } \begin{bmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \mapsto -1.$$

(In the second matrix, only the first 1 is replaced with −1.)

Definition 3.7. The *special orthogonal group*, denoted SO_n, is the kernel of the map $\det\colon \mathrm{O}_n \to \{\pm 1\}$.

Note that the index of SO_n in O_n is 2 — the subgroup SO_n has two cosets in O_n.

§3.1.1 Orthogonal Matrices in Two Dimensions

Question 3.8. What are the matrices in O_2?

To write down an orthogonal matrix, we can write down two vectors which form an orthonormal basis (and take these two vectors as our columns). We can write $\vec{v_1} = (\cos\theta, \sin\theta)^\intercal$ for some angle θ, since $\vec{v_1}$ must have length 1. Then $\vec{v_2}$ must be perpendicular to $\vec{v_1}$ (and must also have length 1), so it must be either $(-\sin\theta, \cos\theta)$ or $(\sin\theta, -\cos\theta)$. In the first case, we get the matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

which has determinant 1; this corresponds to rotation by θ around the origin, which clearly preserves distances. In the second case, meanwhile, we get the matrix

$$\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

Call this matrix A.

Proposition 3.9

The matrix A corresponds to reflection over a line through the origin.

Proof. First, its characteristic polynomial is

$$p_A(t) = \det\begin{bmatrix} t-\cos\theta & -\sin\theta \\ -\sin\theta & t+\cos\theta \end{bmatrix} = t^2 - 1 = (t-1)(t+1).$$

This means we can find an eigenbasis consisting of a vector $\vec{v_+}$ with eigenvalue $+1$, and a vector $\vec{v_-}$ with eigevalue -1. Then we want to show that A is a reflection over $\text{Span}(\vec{v_+})$ (since all vectors on this line are fixed). Note that since A preserves the dot product, we have

$$A\vec{v_+} \cdot A\vec{v_-} = \vec{v_+} \cdot \vec{v_-}.$$

But we also have $A\vec{v_+} = \vec{v_+}$ and $A\vec{v_-} = -\vec{v_-}$, so then

$$A\vec{v_+} \cdot A\vec{v_-} = -\vec{v_+} \cdot \vec{v_-}.$$

This means we must have $\vec{v_+} \cdot \vec{v_-} = 0$, and therefore $\vec{v_+}$ and $\vec{v_-}$ are perpendicular.

So then we have a line $\text{Span}(\vec{v_+})$ which is fixed, and a perpendicular vector $\vec{v_-}$ which is reflected across this line. So if we write any vector as a linear combination of $\vec{v_+}$ and $\vec{v_-}$, it is sent exactly to its reflection over the line. □

In fact, we can compute where the line is (by finding $\vec{v_+}$ and $\vec{v_-}$ explicitly) — it's the line at an angle of $\frac{\theta}{2}$.

> **Remark 3.10.** Note that two reflections across different lines through the origin form a rotation about the origin. We can think of this algebraically (the determinants multiply, and $(-1)(-1) = 1$) or geometrically.

§3.1.2 Orthogonal Matrices in Three Dimensions

So far, we've found a full description for O_n. The situation in three dimensions is a bit more complicated, but it's still possible to describe all of O_3 explicitly. We'll start by answering a slightly simpler question:

> **Question 3.11.** What are the matrices in SO_3?

The answer will still turn out to be rotation matrices. To describe a rotation in three dimensions, we can fix a unit vector $\vec{u} \in \mathbb{R}^3$ and an angle θ, and let $\rho(\vec{u}, \theta)$ be the 3×3 matrix which rotates around the axis $\vec{u}$ by an angle θ. More precisely, $\rho(\vec{u}) = \vec{u}$, and $\rho(\vec{u}, \theta)$ restricted to $\vec{u}^\perp$ (the set of vectors perpendicular to $\vec{u}$) is the matrix corresponding to rotation by θ counterclockwise (where when we say counterclockwise, we're looking in the direction that $\vec{u}$ is sticking out of). This uniquely determines the linear transformation, since every vector can be written as a linear combination of $\vec{u}$ and an element of $\vec{u}^\perp$.

> **Theorem 3.12**
>
> The group SO_3 consists of exactly the matrices $\rho(\vec{u}, \theta)$.

Proof. First we'll show that all matrices $\rho(\vec{u}, \theta)$ are in SO_3. This is unsurprising if we think geometrically, since rotations preserve distance.

Choose vectors $\vec{v}$ and $\vec{w}$ which form an orthonormal basis for the plane $\vec{u}^\perp$, so then $\vec{u}$, $\vec{v}$, and $\vec{w}$ form an orthonormal basis for $\mathbb{R}^3$. We can then create the change of basis matrix

$$P = \begin{bmatrix} | & | & | \\ \vec{u} & \vec{v} & \vec{w} \\ | & | & | \end{bmatrix} \in \text{O}_3.$$

Now the transformation written in the basis $\vec{u}$, $\vec{v}$, $\vec{w}$ is $P^{-1}\rho(\vec{u}, \theta)P$. But this transformation is easy to describe — it fixes $\vec{u}$, and in the plane $\vec{u}^\perp$ it is just a rotation. So we have

$$P^{-1}\rho(\vec{u}, \theta)P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}.$$

It's clear that the matrix on the right is in SO_3; then $P^{-1}\rho(\vec{u},\theta)P$ must be in O_3 (by closure, since P is in O_3) and therefore in SO_3 as well (since the determinants of P and P^{-1} cancel out, so the determinant of $\rho(\vec{u},\theta)$ is the same as the determinant of the right-hand side matrix).

Now we'll show that every matrix in SO_3 must be a rotation matrix $\rho(\vec{u},\theta)$ for some $\vec{u}$ and θ. The first step is to find the axis:

> **Claim** — There is a unit vector $\vec{u}$ with eigenvalue 1.

Proof. It's enough to show that 1 is an eigenvalue of A, or equivalently that $\det(A-I)=0$. But we have

$$\det(A-I)=\det(A^\intercal)\det(A-I)=\det(A^\intercal A-A^\intercal)=\det(I-A^\intercal)=\det(I-A),$$

using the fact that $A^\intercal A=I$ and the determinant of a matrix's transpose is the same as the determinant of the original matrix. But we have $\det(I-A)=(-1)^3\det(A-I)$, so then we must have $\det(A-I)=0$ (since it is equal to its negative).

This means 1 is a root of the characteristic polynomial $p_A(t)$, so it must be an eigenvalue, and it therefore has an eigenvector $\vec{u}$; by scaling, we may assume $\vec{u}$ is a unit vector. ■

Now we're mostly done. We can extend $\vec{u}$ to an orthonormal basis of $\mathbb{R}^3$ by taking an orthonormal basis $\{\vec{v},\vec{w}\}$ of $\vec{u}^\perp$. This again gives us an orthogonal change of basis matrix

$$P=\begin{bmatrix} | & | & | \\ \vec{u} & \vec{v} & \vec{w} \\ | & | & | \end{bmatrix}\in \text{O}_3.$$

We'd like to rewrite the transformation in this basis again, meaning we want to describe the matrix $P^{-1}AP$. This matrix is again in SO_3, since $A\in\text{SO}_3$ and $P\in\text{O}_3$. We know that A sends $\vec{u}$ to itself, so the first column of this matrix must be $(1,0,0)^\intercal$; then since $(1,0,0)^\intercal$ must be orthogonal to both of the other columns, their first entries must be 0. So then our matrix is of the form

$$P^{-1}AP=\begin{bmatrix} 1 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix}.$$

But then the remaining 2×2 matrix must also be orthogonal and have determinant 1, meaning that it must be in SO_2. Using our description of SO_2 from earlier, this means we have

$$P^{-1}AP=\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

for some θ. So this means we have $A=\rho(\vec{u},\theta)$ for these values of $\vec{u}$ and θ. □

We've now described all of SO_3. To describe all of O_3, note that SO_3 has index 2 in O_3, so it's enough to find its other coset. To do so, we can take the reflection matrix

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

which has determinant 1; then the rest of O_3 is the right coset of SO_3 corresponding to this matrix.

§3.2 Isometries

Definition 3.13. A function $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is an *isometry* if $|f(\vec{u}) - f(\vec{v})| = |\vec{u} - \vec{v}|$ for all $\vec{u}, \vec{v} \in \mathbb{R}^n$.

In other words, an isometry is a function which preserves distance. Note that an isometry does *not* have to be a linear transformation, so these are more general than orthogonal matrices.

There's a few obvious examples of isometries — the maps corresponding to orthogonal matrices, meaning $\vec{x} \mapsto A\vec{x}$ for some $A \in \mathrm{O}_n$, are all isometries. All translations $\vec{x} \mapsto \vec{x} + \vec{b}$ for some fixed $\vec{b}$ are also isometries; note that translations are *not* linear transformations, since they don't fix the origin.

Notation 3.14. We use $t_{\vec{b}}$ to denote the translation $\vec{x} \mapsto \vec{x} + \vec{b}$.

Amazingly, it turns out that these are essentially the only ones! Even though an isometry is defined in a much looser way than orthogonal matrices — the function isn't required to be linear — there aren't many new possibilities we get.

Theorem 3.15

All isometries f are a composition $t_{\vec{b}} \circ A$ for some $A \in \mathrm{O}_n$ and $\vec{b} \in \mathbb{R}^n$.

So in other words, every isometry can be written as $f(\vec{x}) = A\vec{x} + \vec{b}$, where A is orthogonal.

To prove this, we'll first consider isometries which fix the origin.

Lemma 3.16

If f is an isometry which fixes the origin, then f must be a linear operator.

Then any isometry which fixes the origin must come from an orthogonal matrix — by taking $\vec{v} = 0$ in the definition, we get that $|f(\vec{u})| = |\vec{u}|$ for all $\vec{u}$, and the linear operators which preserve lengths are exactly the orthogonal matrices.

Proof. First we'll show that f behaves well with respect to the dot product. We can write the dot product in terms of distances between two vectors and 0, as

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

(This can be shown by expanding the right-hand side using $|\vec{x}|^2 = \vec{x} \cdot \vec{x}$.) But since f preserves distances and $f(0) = 0$, if we replace $\vec{u}$ and $\vec{w}$ with $f(\vec{u})$ and $f(\vec{w})$, then the right-hand side is preserved, and we get

$$f(\vec{u}) \cdot f(\vec{v}) = \vec{u} \cdot \vec{v}.$$

Now to show linearity, we can express addition using the dot product: we have $\vec{z} = \vec{x} + \vec{y}$ if and only if

$$(\vec{z} - \vec{x} - \vec{y}) \cdot (\vec{z} - \vec{x} - \vec{y}) = 0,$$

which we can expand out to

$$\vec{z} \cdot \vec{z} + \vec{x} \cdot \vec{x} + \vec{y} \cdot \vec{y} - 2\vec{x} \cdot \vec{z} - 2\vec{y} \cdot \vec{z} + 2\vec{x} \cdot \vec{y} = 0.$$

But since f preserves dot products, this condition is satisfied for $\vec{x}$, $\vec{y}$, and $\vec{z}$ if and only if it's satisfied for $f(\vec{x})$, $f(\vec{y})$, and $f(\vec{z})$; so we have $\vec{z} = \vec{x} + \vec{y}$ if and only if $f(\vec{z}) = f(\vec{x}) + f(\vec{y})$. So then

$$f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}).$$

We can perform the same argument to show that f works well with scalar multiplication — we have $\vec{y} = c\vec{x}$ if and only if $(\vec{y} - c\vec{x}) \cdot (\vec{y} - c\vec{x}) = 0$, and we can expand this out in the same way as before to get that $\vec{y} = c\vec{x}$ if and only if $f(\vec{y}) = cf(\vec{x})$, and therefore $f(c\vec{x}) = cf(\vec{x})$.

So then f must be a linear operator. □

Now the general case is straightforward:

Proof of Theorem 3.15. Given an isometry f, let $\vec{b} = f(0)$, and consider $t_{-\vec{b}} \circ f$. This is also an isometry, and it fixes the origin. So by the above lemma we must have $t_{\vec{b}} \circ f = A$ for some $A \in \mathrm{O}_n$, which means $f = t_{\vec{b}} \circ A$ for some such A. □

So now we have a classification of all isometries. Note that the isometries form a *group* — every isometry has an inverse, which is also an isometry (since translations and orthogonal matrices are both invertible).

Notation 3.17. We use M_n to denote the group of isometries of $\mathbb{R}^n$.

We can think of M_n as a subgroup of the group of *all* permutations of $\mathbb{R}^n$ (bijections from $\mathbb{R}^n$ to itself).

Then the translations form a subgroup of M_n. We can think of this subgroup as just $\mathbb{R}^n$ under addition, since $t_{\vec{b}} \circ t_{\vec{b}'} = t_{\vec{b}+\vec{b}'}$. Meanwhile, the orthogonal matrices also form a subgroup $\mathrm{O}_n \leq M_n$. Theorem 3.15 then says that M_n is generated by these two subgroups.

The theorem writes isometries in the form $t_{\vec{b}} \circ A$. It'll often be useful to convert isometries written in the *opposite* order to this form. In order to simplify $A \circ t_{\vec{b}}$, note that for any $\vec{x}$,

$$A \circ t_{\vec{b}}(\vec{x}) = A(\vec{x} + \vec{b}) = A\vec{x} + A\vec{b},$$

which means that

$$A \circ t_{\vec{b}} = t_{A\vec{b}} \circ A. \tag{1}$$

It'll sometimes be useful to just focus on the orthogonal matrix part of an isometry, and ignore the translation (the constant term):

Notation 3.18. We use π to denote the homomorphism $M_n \to O_n$ sending $t_{\vec{b}} \circ A \mapsto A$.

To see that this is a group homomorphism, note that

$$(t_{\vec{b}} \circ A) \circ (t_{\vec{b}'} \circ A') = t_{\vec{b}} \circ (t_{A\vec{b}'} \circ A) \circ A' = t_{\vec{b}+A\vec{b}'} \circ AA'$$

by applying (1). So if π maps two isometries to A and A', then it maps their product to AA', and therefore π is a homomorphism.

Note that π is surjective, since it maps each orthogonal matrix to itself. Meanwhile, $\ker(\pi)$ is the set of translations. In particular, this means the translations form a normal subgroup of M_n. (It's also possible to see this by using (1), which implies that for any translation t, its conjugate $A \circ t \circ A^{-1}$ is also a translation).

§3.2.1 Isometries in Two Dimensions

Question 3.19. What do the isometries of $\mathbb{R}^2$ look like?

Definition 3.20. An isometry $t_{\vec{b}} \circ A$ is *orientation-preserving* if $\det(A) = 1$, and *orientation-reversing* if $\det(A) = -1$.

So in two dimensions, an isometry is orientation-preserving if A is a rotation matrix, and orientation-reversing if A is a reflection matrix.

Similarly to our description of orthogonal matrices in two dimensions, it's possible to describe *all* isometries in two dimensions as well.

Theorem 3.21

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

(1) A translation;

(2) A rotation about any point $\vec{p}$;

(3) A reflection across a line ℓ (which does not necessarily pass through the origin);

(4) A *glide reflection*, where we reflect across a line ℓ and then translate by a vector $\vec{v}$ parallel to ℓ.

Note that (1) and (2) are orientation-preserving, and (3) and (4) are orientation-reversing.

Proof. The main idea is that given an isometry f, we can shift the origin — the isometry $t_{\vec{p}} \circ f \circ t_{-\vec{p}}$ is the same isometry, but with the origin shifted to $\vec{p}$ (for instance, if f fixed the origin, then this new isometry would fix $\vec{p}$). So we'd like to choose some $\vec{p}$ for which our isometry becomes nicer.

Let f be the isometry $f(\vec{x}) = A\vec{x} + \vec{b}$. We can then use our classification of O_2 from earlier:

Case 1 ($A = I$). Then f is just a translation by $\vec{b}$, corresponding to (1).

Case 2 (A is a non-identity rotation matrix). Then we'd like to find a point $\vec{p}$ fixed by the isometry. First, we know that 1 is not an eigenvalue of A (since rotation matrices don't fix any vectors), so the kernel of $A - I$ is trivial. But then this measn $A - I$ is invertible, and therefore there is a unique solution for $\vec{p}$ to

$$(A - I)\vec{p} = -\vec{b},$$

which rearranges to $f(\vec{p}) = \vec{p}$. Then we can shift $\vec{p}$ to the origin — we can write $f = t_{\vec{p}} \circ g \circ t_{-\vec{p}}$ for an isometry g fixing the origin. Then g must be a rotation about the origin; so f is a rotation about $\vec{p}$.

Case 3 (A is a reflection matrix). We can again use a similar idea of shifting the origin — first write $f = t_{\vec{b}} \circ A$, where A corresponds to reflection across some line ℓ through the origin. Now we shift the origin by $\frac{1}{2}\vec{b}$ — consider the isometry

$$g = t_{-\vec{b}/2} \circ f \circ t_{\vec{b}/2} = t_{\vec{b}/2} \circ A \circ t_{\vec{b}/2} = t_{\vec{b}/2} \circ t_{A\vec{b}/2} \circ A = t_{\vec{m}} \circ A,$$

where $\vec{m} = \frac{1}{2}(\vec{b} + A\vec{b})$. Note that $\vec{m}$ is the average of $\vec{b}$ and its reflection over ℓ, so $\vec{m}$ is necessarily parallel to ℓ. If $\vec{m} = 0$ then g is a reflection about ℓ, while otherwise g is a glide reflection about ℓ. Then $f = t_{\vec{b}/2} \circ g \circ t_{\vec{b}/2}$ is the same isometry with the origin shifted to $\frac{1}{2}\vec{b}$ — a reflection or glide reflection about ℓ shifted by $\frac{1}{2}\vec{b}$. □

§3.3 What Is Symmetry?

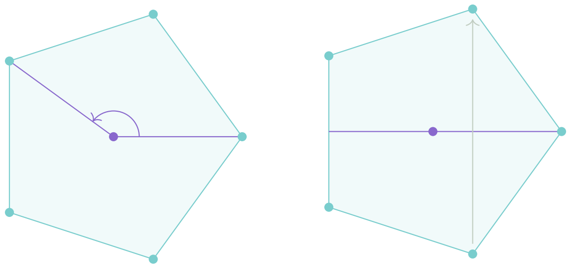
Question 3.22. What isometries of $\mathbb{R}^2$ fix some pattern in $\mathbb{R}^2$?

Definition 3.23. Given a figure P, the *symmetries* of P are the isometries that fix P.

Example 3.24

What are the symmetries of a regular pentagon? What about a circle?

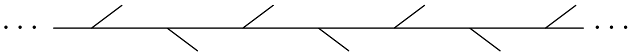
Solution. For the pentagon, we can rotate by $2\pi k/5$ for any integer k, or reflect across any line through one vertex and the center:



Meanwhile, for the circle, all rotations and reflections work. □

Note that the symmetries of a pentagon are *discrete*, and the symmetries of a circle are not. (We will make this notion more precise later.)

We can also consider the symmetries of an infinite figure. For example, an equilateral triangular lattice has symmetries of each kind — translations, reflections, rotations, and glide reflections. Meanwhile the symmetries of the following shape are translations and glide reflections:



In both cases, the set of symmetries is infinite, but still discrete.

Question 3.25. What kinds of subgroups of M_2 can we get in this way?

To make this question more precise, we'll try to describe all *discrete* subgroups of M_2.

§3.4 Discrete Subgroups of O_2

First we'll first look at discrete subgroups of O_2. To begin with, we'll define discrete subgroups in a simpler setting, the real numbers.

Definition 3.26. A subgroup $G \leq (\mathbb{R}, +)$ is *discrete* if there exists $\varepsilon > 0$ such that for all nonzero $g \in G$, we have $|g| > \varepsilon$.

This is equivalent to requiring that no two points in G can be too close together — for any two elements a and b with $a \neq b$, we must have $|a - b| > \varepsilon$ (since $a - b$ is in G).

Theorem 3.27

If $G \leq (\mathbb{R}, +)$ is discrete, then either $G = \{0\}$ or $G = \mathbb{Z}\alpha$ for some $\alpha > 0$.

Proof. Assume G is not zero. Then we claim it has a smallest positive element α — take any positive element $g > 0$ in G. Then since every two elements are a distance at least ε apart, there are finitely many elements in $[0, g)$, and therefore one of them is the smallest nonzero element.

Then since G is a group, it contains $n\alpha$ for every integer α. We claim that there are no other elements — assume there is some x with $n\alpha < x < (n+1)\alpha$. Then we have $0 < x - n\alpha < \alpha$. But $x - n\alpha$ must be in G, contradicting the choice of α as the smallest element. □

Remark 3.28. This is very similar to our proof that the only subgroups of $\mathbb{Z}$ are $\{0\}$ and $n\mathbb{Z}$ for positive integers n — the only difference here is that here we used discreteness in order to prove that there exists a smallest element.

We'll start with a slightly simpler questoin:

Question 3.29. What are the *finite* subgroups of O_2?

There's a few obvious examples. If x is a rotation by $2\pi/n$, then the cyclic group $C_n = \langle x \rangle = \{1, x, \dots, x^{n-1}\}$ is a finite subgroup of O_2; in fact, it's a finite subgroup of SO_2 as well.

We can also let y be the reflection across some line ℓ through the origin, and consider the group $\langle x, y \rangle$. We have the relations $yx = x^{-1}y$ and $x^n = y^2 = e$, so

$$D_n = \langle x, y \rangle = \{e, x, x^2, \dots, x^{n-1}, xy, x^2y, \dots, x^{n-1}y\}$$

is also a finite subgroup of O_2 (but not SO_2).

Definition 3.30. The group D_n (generated by a rotation by $2\pi/n$ and a reflection) is called the *dihedral group*.

For example, we have $D_1 \cong C_2$, $D_2 \cong C_2 \times C_2$, and $D_3 \cong S_3$. For $n \geq 3$, the dihedral group D_n is the group of symmetries of a regular n-gon. Note that C_n is always a subgroup of D_n with index 2.

So we've seen a few finite subgroups of O_2. It turns out these are the only ones!

Theorem 3.31

Every finite subgroup of O_2 is isomorphic to C_n or D_n for some n.

To prove this, we'll first prove a more specific case.

Lemma 3.32

Every finite subgroup of SO_2 is isomorphic to C_n for some n.

Proof. We know that SO_2 consists exactly of the rotation matrices

$$\rho_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Let our subgroup be H, and let $S = \{\theta \in \mathbb{R} \mid \rho_\theta \in H\}$ be the set of angles which appear in H. Then since H is a group, S must be a subgroup of $\mathbb{R}$; and since H is finite, S must be discrete. So by Theorem 3.27, S must be of the form $\mathbb{Z}\alpha$ for some $\alpha > 0$. But we also know $2\pi \in S$, since rotation by 2π is the identity; therefore we must have $2\pi = n\alpha$ for some positive integer n, and therefore H is generated by a rotation by $2\pi/n$. □

Now using this, we can prove the theorem for all subgroups $G \leq \mathrm{O}_2$.

Proof of Theorem 3.31. First, if $G \leq \mathrm{SO}_2$, then the above lemma implies that $G \cong C_n$ for some n. So now assume G is not contained in SO_2, so G has a reflection, or equivalently an element of determinant -1.

Now consider the determinant, which gives a homomorphism $\det\colon G \to \{\pm 1\}$. Then since G has elements of determinant -1, this homomorphism must be surjective. So its kernel H is a normal subgroup of G with index 2, and its two cosets are H itself and Hr for some reflection r.

But since H is a finite subgroup of SO_2, we must then have $H \cong C_n$ for some n. Then we have $G = \langle r\rho_{2\pi/n}, r\rangle$, which means $G \cong D_n$. □

So we've classified all finite subgroups of O_2. We really wanted to classify all *discrete* subgroups. First we need a more precise definition of a discrete subgroup of O_2:

> **Definition 3.33.** A subgroup $G \leq \mathrm{O}_2$ is discrete if there exists some $\varepsilon > 0$ such that all rotations ρ_θ in G have $|\theta| > \varepsilon$.

Now with this definition, the same argument as the one we used in the finite case works; in particular, the conclusion is the same.

§3.5 Discrete Subgroups of Isometries

Now that we've understood the discrete subgroups of O_2, we can try to understand the discrete subgroups of M_2 (the group of isometries).

> **Definition 3.34.** A subgroup $G \leq M_2$ is discrete if there exists $\varepsilon > 0$ such that all translations $t_{\vec{b}}$ in G have $|\vec{b}| > \varepsilon$, and all rotations ρ_θ in G have $|\theta| > \varepsilon$.

Intuitively, this means the angles of rotation and the translations must both be discrete. So we're avoiding groups like the symmetries of a circle (where we had *all* rotations, for example), but the symmetries of most "reasonable" shapes (such as a triangular lattice) are discrete.

As we'll see later, this ends up being quite a strong constraint.

§3.5.1 Finite Subgroups of Isometries

As with the case of O_2, we'll start by asking a simpler question:

> **Question 3.35.** What are the *finite* subgroups of M_2?

Of course, all the finite subgroups of O_2 are still finite subgroups of M_2. It turns out that these are still the only ones.

> **Theorem 3.36**
>
> Every finite subgroup of M_2 is isomorphic to C_n or D_n for some n.

Proof. Let the group be $G \leq M_2$. The main idea is to find a point fixed by all isometries in G; then we can shift coordinates so that this fixed point is the origin, which reduces the question to finite subgroups of O_2.

Claim — Any finite group of isometries has a fixed point.

Proof. First we'll find a finite *set* S which is preserved by G, meaning that $g(S) = S$. Fix any point $p \in \mathbb{R}^2$, and take the set

$$S = \{g(p) \mid g \in G\}.$$

As we'll see later, this set is called the *orbit* of p. Then S is finite since G is finite. But if we take a point $s = h(p) \in S$ and apply g to s, then we get

$$g(s) = g(h(p)) = (gh)(p) \in S$$

as well, since G is closed. (Technically we've only shown that g is a map from S to itself; but this map must actually be a bijection, since g has an inverse, so $g(S)$ is genuinely equal to S.)

Now we've found a finite *set* fixed by G, and we want to find a single *point* fixed by G. To do so, we can simply take the *average* of all points in S — take

$$s_0 = \frac{1}{n}(s_1 + \cdots + s_n).$$

Isometries play well with averages — if $f = t_{\vec{b}} \circ A$ is some isometry, then we have

$$f(s_0) = \vec{b} + \frac{1}{n}(As_1 + \cdots + As_n) = \frac{1}{n}\sum_{i=1}^{n}(\vec{b} + As_i) = \frac{1}{n}(f(s_1) + \cdots + f(s_n)).$$

So for each $g \in G$, we have that $g(s_0)$ is the average of all points in $g(S)$. But since $g(S) = S$, this means $g(s_0)$ is also the average of all points in S, which is just s_0. So s_0 is a fixed point of all $g \in G$. ■

Now we can shift our coordinate system so that the fixed point s_0 is the origin. Then all our isometries correspond to orthogonal matrices, which means $G \leq \mathrm{O}_2$, and therefore G must be C_n or D_n for some n by Theorem 3.31. □

Remark 3.37. The reason finiteness was needed here is so that we could take the average.

Unlike the case of O_2, though, it turns out that the *discrete* subgroups of M_2 are more complicated.

§3.5.2 Discrete Subgroups of $\mathbb{R}^2$

We've already looked at the discrete subgroups of O_2. Another fairly simple subgroup sitting inside M_2 is the group of translations, which we can think of as $\mathbb{R}^2$. So we can try to analyze the discrete subgroups of these translations first.

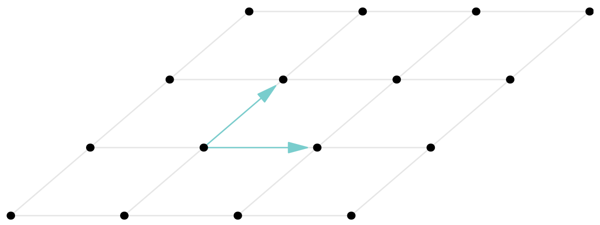
Question 3.38. What are the discrete subgroups of $\mathbb{R}^2$?

We've already answered this question for $\mathbb{R}$ — all discrete subgroups of $\mathbb{R}$ are either $\{0\}$ or $\mathbb{Z}\alpha$ for some $\alpha > 0$. The answer for $\mathbb{R}^2$ turns out to be fairly similar.

Theorem 3.39

If G is a discrete subgroup of $\mathbb{R}^2$, then G is either $\{0\}$, the group $\mathbb{Z}\alpha$ for some vector α, or the group $\mathbb{Z}\alpha + \mathbb{Z}\beta$ for some linearly independent vectors α and β.

In the third case, G is called a *lattice*; it looks like a parallelogram grid.



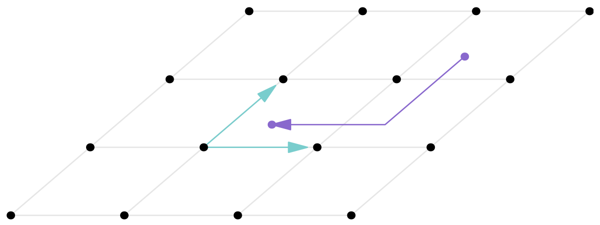
Proof. The proof is very similar to the one-dimensional case. Assume G is not 0, and pick some nonzero element $\widetilde{\alpha}$ in G. First we consider $G \cap \mathbb{R}\widetilde{\alpha}$, the elements of G which lie on the line spanned by $\widetilde{\alpha}$. This is a discrete subgroup of $\mathbb{R}\widetilde{\alpha}$, so by the one-dimensional case (in Theorem 3.27), it must be of the form $\mathbb{Z}\alpha$ for some α.

Now if there are no other vectors in G, we're done. So assume that G does contain other vectors.

> **Claim** — There exists a vector $\beta \notin \mathbb{R}\alpha$ with minimal distance to the line $\mathbb{R}\alpha$.

Proof. First, in any bounded region, there can only be finitely many elements of G (since all elements of G are at distance greater than ε from each other, so we can tile any bounded region with finitely many balls which can each contain at most one point).

Now take any $\widetilde{\beta}$ not on the line $\mathbb{R}\alpha$. Then it suffices to consider points inside the parallelograms with sides α and $\widetilde{\beta}$ and with sides α and $-\widetilde{\beta}$ — any point can be brought inside each parallelogram by subtracting multiples of α and $\widetilde{\beta}$, and one of these new points is at least as close to $\mathbb{R}\alpha$ as the original point.



But since these parallelograms are bounded, there are finitely many points inside them, so we can pick some β which is closest to the line. ■

Now let β be such a point; then we claim $G = \mathbb{Z}\alpha + \mathbb{Z}\beta$. Assume not, so G contains some point not in $\mathbb{Z}\alpha + \mathbb{Z}\beta$. Then we can shift by α and β to bring this point inside the parallelogram with sides α and β. Either this point is on $\mathbb{R}\alpha$ — contradicting the choice of α as the element of the line with smallest magnitude — or it's strictly closer to $\mathbb{R}\alpha$ than β is, contradiction. □

§3.5.3 The Point Group

Now we've studied discrete subgroups of both O_2 and $\mathbb{R}^2$, and we can return to our general question:

> **Question 3.40.** How do we study the discrete subgroups of M_2?

Recall that in some sense, all elements of M_2 can be built from O_2 and $\mathbb{R}^2$. More precisely, as mentioned earlier, we have a surjective homomorphism $\pi\colon M_2 \to \mathrm{O}_2$ which ignores translations and just keeps track of the linear term (so if $f(\vec{x}) = A\vec{x} + \vec{b}$, then π sends $f \mapsto A$), and $\ker(\pi)$ is precisely the subgroup of translations.

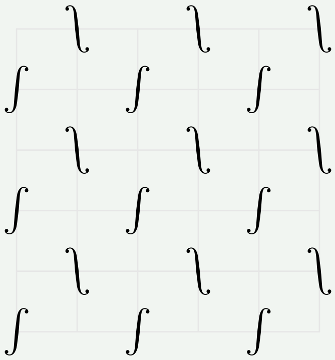
Since $G \leq M_2$, we can restrict π to G. Then the image of π is a subgroup of O_2 consisting of the linear parts of elements in G, called the *point group* of G — so the point group keeps track of the angle of rotation or the slope of the line used for reflection, but it *doesn't* keep track of any gliding, or the location of the point or line. Meanwhile, the kernel of π consists of exactly the translations inside G.

> **Notation 3.41.** We use $\widetilde{G}$ to denote the point group of G (or equivalently, the image of π restricted to G), and L to denote the kernel of π restricted to G.

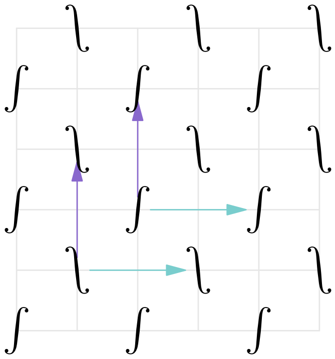
Then since G is a discrete subgroup of M_2, the point group $\widetilde{G}$ must be a discrete subgroup of O_2, and we've already solved what these are! Meanwhile L must also be a discrete subgroup of $\mathbb{R}^2$, and we've solved what these are as well. So there's two possibilities for $\widetilde{G}$ — it can be C_n or D_n for any n — and three for L — it can be $\{0\}$, $\mathbb{Z}\alpha$, or $\mathbb{Z}\alpha + \mathbb{Z}\beta$ for some vectors α and β.

Example 3.42

Find L and $\widetilde{G}$ for the following (infinite) picture:

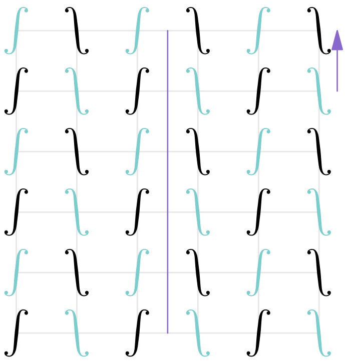


Solution. First, L is a lattice — we can translate both horizontally and vertically.



Now we'll find $\widetilde{G}$. First, G contains rotation by 180° about the center of any of the integrals, so $\widetilde{G}$ contains a rotation by 180° (since $\widetilde{G}$ doesn't keep track of the point we're rotating about).

Meanwhile, G also contains a glide reflection — we can reflect about a line between two columns, and then glide vertically to align the reflected picture with the initial one.

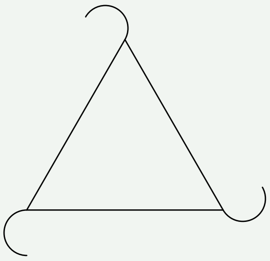


So then $\widetilde{G}$ contains a reflection (the point group doesn't keep track of any gliding).

This means $\widetilde{G}$ contains a rotation by 180° and a reflection, so $\widetilde{G} = D_2$. □

Example 3.43

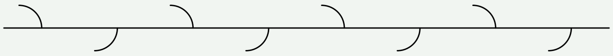
Find L and $\widetilde{G}$ for the following picture:



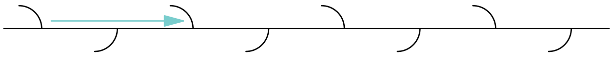
Solution. First, no nontrivial translations preserve the shape, so $L = \{0\}$. Meanwhile we can rotate by 120°, but we can't reflect (since reflections change the orientation of the semicircles), so $\widetilde{G} = C_3$. □

Example 3.44

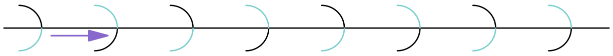
Find L and $\widetilde{G}$ for the following (infinite) picture:



Solution. First, we can translate horizontally, so $L = \mathbb{Z}\alpha$ where α is the vector between two arcs on the same side of the line:



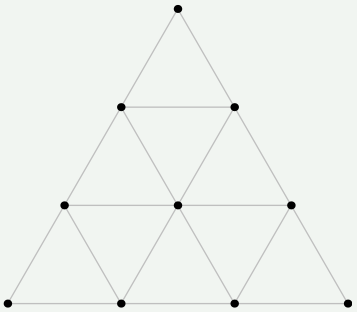
This figure doesn't have any rotational symmetries. It doesn't have any reflections either, but it does have a *glide* reflection:



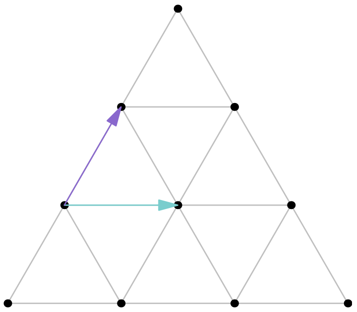
□

Example 3.45

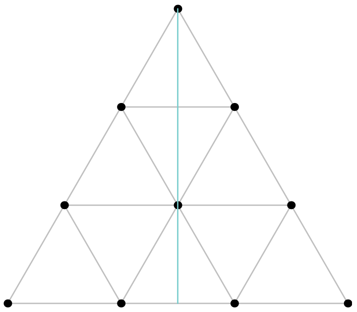
Find L and $\widetilde{G}$ for an (infinite) equilateral triangular grid:



Solution. First, the translations form a lattice:



Meanwhile, we can rotate by 60° about one of the points in the lattice. We can also reflect:



So then we have $\widetilde{G} = D_6$. □

§3.5.4 Crystallographic Restriction

So far, we've described the possibilities for $\widetilde{G}$ and L separately. But it turns out that we can get a lot more information by looking at how they interact with each other.

Theorem 3.46

The point group $\widetilde{G}$ must map L to itself.

In other words, if $A \in \widetilde{G}$ and $\vec{b} \in L$, then we must have $A\vec{b} \in L$ as well.

Proof. Since A is in $\widetilde{G}$, there is some vector $\vec{c}$ for which $t_{\vec{c}} \circ A$ is in G. Meanwhile since $\vec{b}$ is in L, it must also be in G.

But L is the kernel of the homomorphism $\pi\colon G \to \widetilde{G}$, so L is a normal subgroup of G. Then L is preserved under conjugation by any element of G, so in particular we have

$$(t_{\vec{c}} \circ A) \circ t_{\vec{b}} \circ (t_{\vec{c}} \circ A)^{-1} \in L.$$

But we can now expand this out as

$$t_{\vec{c}} \circ A \circ t_{\vec{b}} \circ A^{-1} \circ t_{\vec{c}}^{-1} = t_{\vec{c}} \circ t_{A\vec{b}} \circ AA^{-1} \circ t_{-\vec{c}} = t_{\vec{c}} \circ t_{A\vec{b}} \circ t_{-\vec{c}} = t_{A\vec{b}}.$$

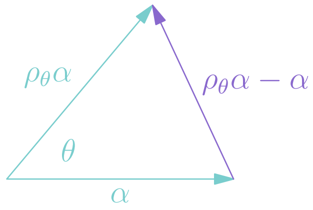
So we must have $A\vec{b} \in L$, as claimed. □

This is a very strong constraint! Given a group $\widetilde{G}$, most possible lattices won't be preserved by it — knowing that a lattice is preserved tells us something special about its angles.

Theorem 3.47 (Crystallographic Restriction)
If L is nonzero, then $\widetilde{G}$ must be C_n or D_n for some $n \in \{1, 2, 3, 4, 6\}$.

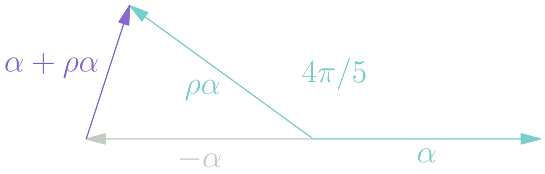
Proof. Let α be a nonzero vector in L of minimal length (which exists because L is discrete), and suppose we have a rotation $\rho_\theta \in \widetilde{G}$. Then since the rotation must preserve L, then we must have $\rho_\theta\alpha \in L$ as well.

But since L is a lattice, their difference $\rho_\theta\alpha - \alpha$ must be in L as well. But if $\theta < \pi/3$, then this vector is strictly shorter than α, contradiction.



So $\widetilde{G}$ cannot contain any rotations by $\theta < \pi/3$. Since we know $\widetilde{G}$ must be C_n or D_n for *some* n, this means we must have $n \leq 6$.

It now remains to eliminate the case of $n = 5$. Let ρ be the rotation by $4\pi/5$, and consider $\alpha + \rho\alpha$. By the same reasoning as before, this vector is again shorter than α.



So then we must have $n \in \{1, 2, 3, 4, 6\}$. □

All such $\widetilde{G}$ are possible — for example, the equilateral triangular lattice in Example 3.45 has point group D_6, and a square lattice has point group D_4.

In fact, given which group $\widetilde{G}$ is, we can constrain L further, and use this to constrain G as well. It turns out that when L is a lattice (meaning G contains two independent translation vectors), there are only 17 possible groups G! (On the other hand, when L is $\{0\}$, there are infinitely many — G can be C_n or D_n for any n.)

Example 3.48

If $\widetilde{G} = C_4$, what can we say about G?

Solution. First, we have a surjective homomorphism $\pi\colon G \to C_4$ with kernel L, so $[G : L] = 4$.

Now let $\widetilde{p} \in \widetilde{G}$ be the rotation by 90°. Then if α is the shortest vector in L, we must also have $\rho\alpha \in L$, and we can show that these two vectors generate L, so $L = \mathbb{Z}\alpha + \mathbb{Z}(\rho\alpha)$.

Then we can take $\rho \in G$ such that $\pi(\rho) = \widetilde{\rho}$. By our classification of isometries in Theorem 3.15, we know ρ is a rotation by 90° about some point; we can choose our coordinate system so that this point is the origin.

Now we can use the fact that the square lattice L has index 4 — its four cosets are given by multiplication by ρ^i for $0 \le i \le 3$, so then

$$G = \{t_{\vec{v}} \circ \rho^i \mid \vec{v} \in L \text{ and } 0 \le i \le 3\}.$$

We also know how to multiply elements of G — we can repeatedly use the fact that

$$\rho \circ t_{\vec{v}} = t_{\rho\vec{v}} \circ \rho.$$

So then G (up to isomorphism) is completely determined from the fact that $\widetilde{G}$ is C_4 (and L is nonzero)! □

In this case, we could completely determine G. The case where $\widetilde{G}$ is D_4 instead of C_4 is more subtle — if $\pi(\rho)$ is a rotation then ρ must also be a rotation, but if $\pi(\rho)$ is a reflection then ρ may be a *glide* reflection instead. In fact, G may not even contain any reflections, as in Example 3.42 (which only has a glide reflection).

But it's still possible to perform a similar analysis. If we take a reflection $\widetilde{r} \in \widetilde{G}$ and take some r with $\pi(r) = \widetilde{r}$, then we can write $r = t_{\vec{b}} \circ r_\ell$ where r_ℓ is a reflection across the line ℓ, and $\vec{b}$ is parallel to ℓ (and may or may not be zero).

We can get additional constraints on $\vec{b}$ — if we compose a glide reflection with itself, then we get

$$t_{\vec{b}} \circ r_\ell \circ t_{\vec{b}} \circ r_\ell = t_{2\vec{b}},$$

using the fact that $\vec{b}$ is parallel to ℓ. So then $2\vec{b}$ must be in the square lattice we obtained for L, and therefore $\vec{b}$ is either in the lattice, or halfway between two of its points.

Remark 3.49. We've essentially seen how to classify all discrete subgroups of isometries in $\mathbb{R}^2$. It's possible to perform a similar analysis for discrete isometries of $\mathbb{R}^3$, but there are a lot more possibilities.

§4 Group Actions

§4.1 Definitions

We've seen many situations where elements of a group are *acting* on some other objects.

Example 4.1

Given a matrix $g \in \mathrm{GL}_n(\mathbb{R})$ and a column vector $\vec{v} \in \mathbb{R}^n$, we can produce a new vector $g\vec{v} \in \mathbb{R}^n$. So we can think of matrix multiplication as a map $\mathrm{GL}_n(\mathbb{R}) \times \mathbb{R}^n \to \mathbb{R}^n$ sending $(g, \vec{v}) \mapsto g\vec{v}$.

Example 4.2

The elements of S_n are permutations of $\{1, 2, \dots, n\}$. So each $\sigma \in S_n$ defines a function on $\{1, 2, \dots, n\}$, and we can think of this as a map $S_n \times \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$ given by $(\sigma, i) \mapsto \sigma(i)$.

Example 4.3

Isometries are functions on $\mathbb{R}^2$, so we can define a map $M_2 \times \mathbb{R}^2 \to \mathbb{R}^2$ as $(f, \vec{x}) \mapsto f(\vec{x})$.

This leads to the concept of a group action:

Definition 4.4. Given a group G and a set S, an *action* of G on S is a map $G \times S \to S$, denoted by $(g, s) \mapsto gs$, which satisfies the following axioms:

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

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

Many of the groups we've seen so far already *come* with an action on some set. But we can take the same group and have it act on *many* different sets at the same time, and we can use this to study the group.

Example 4.5

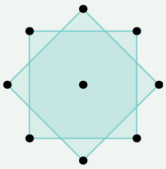
The group S_4 acts on the set $S = \{1, 2, 3, 4\}$ in the obvious way. But it also acts on the set

$$T = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}\},$$

where $\sigma(\{i, j\}) = \{\sigma(i), \sigma(j)\}$.

Example 4.6

The group D_2 acts on the following set S of nine points (by rotating and reflecting the figure):



Example 4.7

Every group G acts on itself, with the action $(g, g') \mapsto g \cdot g'$. (This is an action $G \times G \to G$, where we think of the second copy of G as a set.)

Example 4.8

Given a vector space V over a field F, the group $F^\times$ (the nonzero elements of F under multiplication) acts on V by sending $(a, \vec{v}) \mapsto a\vec{v}$. (The conditions for this to be a group action are a subset of the conditions for V to be a vector space.)

For each $g \in G$, the group action defines a map $\tau_g \colon S \to S$ given by $s \mapsto g(s)$ — so τ_g keeps track of what a fixed element g does to S.

Proposition 4.9

The map τ_g is a bijection.

Proof. The map τ_g has an inverse $\tau_{g^{-1}}$, since for any s we have

$$g(g^{-1}s) = (gg^{-1})(s) = e(s) = s,$$

and similarly $g^{-1}(g(s)) = s$. □

So then $\tau_g \in \text{Perm}(S)$ for all g, which means we can define a map $\tau \colon G \to \text{Perm}(S)$ sending $g \mapsto \tau_g$. The axioms then imply that τ is a group homomorphism — so another way we can think of a group action is as a homomorphism from G to $\text{Perm}(S)$. Note that τ doesn't have to be injective though — it's possible that some $g \in G$ other than the identity fix all elements of S.

§4.2 Orbits and Stabilizers

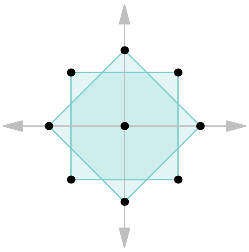
Question 4.10. Given a group action of G on S, what kind of structure do we get?

Definition 4.11. Given an element $s \in S$, the *orbit* of s, denoted O_s or Gs, is the set $\{gs \mid g \in G\}$.

So for each $s \in S$, the orbit of s is a subset of S. Note that s is in its own orbit, since $es = s$.

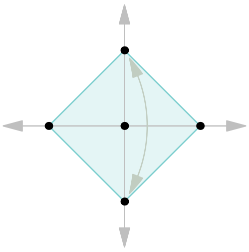
Example 4.12

Consider the action of D_2 on the set S of nine points in Example 4.6, where D_2 consists of rotation by 180°, reflection across the x-axis and y-axis, and the identity. What are the orbits of this action?

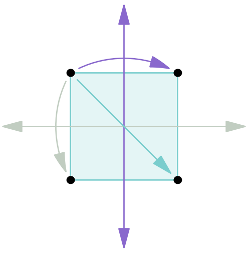


Solution. First, the origin is fixed by all elements of D_2, so it is in its own orbit.

For each of the points on the diamond, the rotation and one of the reflections both send it to the opposite point; meanwhile the other reflection fixes it. So its orbit consists of itself and the point opposite it.



Meanwhile, the square forms an orbit — for any point of the square, the two reflections send it to adjacent points on the square, and the rotation by 180° sends it to the opposite point.



So then we have four orbits — one orbit of size 1, containing the origin; two orbits of size 2, containing pairs of opposite vertices of the diamond; and one orbit consisting of the entire square. □

Note that the orbits don't necessarily have the same size.

Definition 4.13. If there is some $s \in S$ with $O_s = S$, then we say G acts *transitively* on S.

It's possible to check from the axioms that if the orbit of *some* $s \in S$ is the entire set S, then the orbit of *every* s is also S.

Example 4.14

The action of S_n on $\{1, 2, \ldots, n\}$ is transitive, since the orbit of 1 consists of all elements 1, $\ldots$, n (for each k, there is some permutation sending $1 \mapsto k$).

There's also another piece of information we can look at from a group action:

Definition 4.15. Given an element $s \in S$, the *stabilizer* of S, denoted $\operatorname{Stab}(S)$ or G_s, is the set $\{g \mid g(s) = s\}$.

In other words, the stabilizer of s consists of the elements $g \in G$ which fix s. Note that the stabilizer of any s is a *subgroup* of G (this is straightforward to check from the axioms).

Example 4.16

In the group action of D_2 in Example 4.6, the stabilizers are the following:

- The stabilizer of the origin is all of D_2.
- For each point on the diamond, its stabilizer consists of the identity and one reflection.
- The stabilizer of each point on the square is trivial.

Proposition 4.17

The orbits of G form a partition of S.

Proof. First, the orbits cover all elements of S, since each element $s \in S$ is in its own orbit O_s.

Now it remains to show that the orbits are disjoint — meaning that if two orbits have some element in common, they must be the *same* orbit. Suppose the orbits O_s and $O_{s'}$ both have an element t, so we have $t = gs = g's'$ for some g and g' in G. Then

$$s = g^{-1}(t) = g^{-1}(g'(s')) = (g^{-1}g')(s'),$$

so s is in $O_{s'}$. But then $O_s \subset O_{s'}$ as well, since for any $hs \in O_s$ we have

$$hs = h(g^{-1}g(s')) = (hg^{-1}g)s' \in O_s.$$

The same reasoning also shows that $O_{s'} \subset O_s$, so they must be the same set. □

For example, the action of D_2 split the set S of nine points into four orbits, as described in Example 4.12.

Corollary 4.18

If S is finite, then its size is the sum of the sizes of the distinct orbits.

For example, in the action of D_2 we had $9 = 1 + 2 + 2 + 4$.

Question 4.19. What does each orbit O_s look like?

To answer this, we'll look at the stabilizers.

Proposition 4.20

Fix some element $s \in S$, and let $H = \text{Stab}(s)$. Then there is a bijection ε from the left cosets of H to O_s, sending $gH \mapsto gs$.

This has an important corollary:

Corollary 4.21

For each $s \in S$, we have $|O_s| = [G : \text{Stab}(s)]$. In particular, if G is finite then

$$|G| = |\text{Stab}(s)| \cdot |O_s|,$$

and therefore all orbits have size dividing $|G|$.

Proof of Proposition 4.20. First we'll figure out when two elements of G send s to the same element of its orbit. Given two elements g and γ of G, we have $gs = \gamma s$ if and only if $s = g^{-1}\gamma s$. This occurs exactly when $g^{-1}\gamma \in H$, or equivalently when $\gamma \in gH$.

So then g and γ send s to the same element if and only if they're in the same coset of H. This means ε is well-defined and injective.

Meanwhile, ε is surjective as well, since every $s' \in O_s$ is of the form gs for some $g \in G$. So ε is a bijection. □

This corollary is quite useful — as we'll see in the following example, it can let us deduce information about a group by looking at its actions.

Example 4.22

Let $G \le \mathrm{SO}_3$ be the group of rotational symmetries of a cube. Find $|G|$.

Solution. We can consider the action of G on the faces of the cube; let the set of faces be S. Given any two faces, there exists a rotation sending one to the other, so this action is transitive. Since $|S| = 6$, this means there is one orbit of size 6.

Meanwhile, we can also calculate the size of the stabilizers. Consider some face s. For a rotation to fix s, its axis must be perpendicular to s, and its angle must be a multiple of 90°. So $\mathrm{Stab}(s)$ is the cyclic group of order 4.

So then using Corollary 4.21, we have $|G| = 4 \cdot 6 = \boxed{24}$.

We could have performed this argument using vertices or edges instead. For example, let T be the set of vertices, and consider the action of G on T. This action is still transitive, and we now have $|T| = 8$. Meanwhile, for a rotation to preserve a vertex v, its axis must be the long diagonal through v, and there are three possible rotations (since the rotation must preserve the three edges from v). So $\mathrm{Stab}(v) = C_3$, and we get $|G| = 3 \cdot 8 = 24$. □

It's possible to perform similar arguments to find the number of rotational symmetries of other shapes as well, such as a regular tetrahedron or icosahedron.

There's another interesting question to analyze about the structure of orbits and stabilizers:

Question 4.23. How does the stabilizer change across different elements of the same orbit?

Take some element $s' \in O_s$, and suppose $s' = as$ for some $a \in G$. Then for any $g \in \mathrm{Stab}(s)$, we have $gs = s$, which means

$$aga^{-1}(s') = aga^{-1}(as) = ag(s) = as = s'.$$

So if $g \in \mathrm{Stab}(s)$, then $aga^{-1} \in \mathrm{Stab}(s')$, and the converse can be shown similarly. So then the stabilizers of s and s' are *conjugate* — we have

$$\mathrm{Stab}(as) = a\,\mathrm{Stab}(s)a^{-1}.$$

Note that if $\mathrm{Stab}(s)$ is normal, then s and s' have the *same* stabilizer; but $\mathrm{Stab}(s)$ generally does not have to be normal.

§4.3 Finite Subgroups of SO_3

Recall that SO_3 consists of exactly the rotation matrices — every matrix in SO_3 is a rotation around an axis $\vec{u}$ by an angle θ, and this matrix is denoted $\rho(\vec{u}, \theta)$.

Question 4.24. What are the finite subgroups of SO_3?

We've previously answered this question for SO_2. The case of SO_3 is more difficult, but it turns out that with the tool of group actions, we can now answer it here as well.

Theorem 4.25

Every finite subgroup of SO_3 is of one of the following forms:

- The cyclic group C_n, obtained by $\langle \rho(\vec{u}, 2\pi/n)\rangle$ for some $\vec{u}$;
- The dihedral group D_n, obtained by $\langle \rho(\vec{u}, 2\pi/n), \rho(\vec{v}, \pi)\rangle$ for some $\vec{v} \perp \vec{u}$;
- The rotational symmetries of a regular polyhedron — a tetrahedron, cube, octahedron, dodecahedron, or icosahedron.

Note that reflection across some axis in two dimensions corresponds to a 180° rotation about that axis in three dimensions — this is why $\langle \rho(\vec{u}, 2\pi/n), \rho(\vec{v}, \pi)\rangle$ is D_n.

In fact, the last case is somewhat redundant. The cube and octahedron are dual — if we start with a cube and draw the midpoint of each face, this gives an octahedron, and doing the same to an octahedron gives a cube. So any rotational symmetry of the cube gives a rotational symmetry of the octahedron, and vice versa; so their groups of rotational symmetries are the same. Similarly, the rotational symmetries of a dodecahedron and icosahedron are also the same. So there's only three additional subgroups (other than C_n and D_n), and it's possible to analyze these subgroups the same way as we did for a cube in Example 4.22.

Let G be a finite subgroup of SO_3. Then the main idea is to find an action of G, and study its orbits.

Definition 4.26. Given a non-identity element $g \in \mathrm{SO}_3$, its *poles* are the two unit vectors it fixes.

So the poles of a rotation $\rho(\vec{u}, \theta)$ are $\pm\vec{u}$.

Now let P be the set of poles of all the non-identity elements of G.

Lemma 4.27

Our group G acts on P. In other words, for any $p \in P$ and $g \in G$, we have $gp \in P$ as well.

Proof. Suppose p is the pole of some $h \in G$, so then we have $hp = p$. Now let $p' = gp$, so we want to show that p' is also a pole of some element of G. But we have

$$ghg^{-1}(p') = ghp = gp = p'.$$

(Note that this is the same reasoning we used to analyze the stabilizer of $s' = as$.) We know $ghg^{-1} \in G$, and ghg^{-1} cannot be the identity (since h is not the identity). So $p' = gp$ is also in P. □

Example 4.28

When G is C_n, all rotations are about the same axis, so the only poles are p and $-p$ for some point p.

Example 4.29

When G is the group of rotational symmetries of the octahedron (which is denoted O), we have one pole corresponding to each face, vertex, and edge (since each gives a rotation axis).

Now that we have an action, we can analyze its orbits and stabilizers. Let $|G| = n$, and suppose P decomposes into orbits as

$$P = O_1 \cap O_2 \cap \cdots \cap O_k.$$

Let $|O_i| = n_i$ for each i, and let O_i be the orbit of the pole p_i. Finally, let $|\text{Stab}(p_i)| = r_i$, so then we have $n_i r_i = n$ for each i.

Example 4.30

When G is C_n, there are two poles and both are fixed by all elements of G, so we have two one-element orbits, and the stabilizer of each pole is all of G.

Example 4.31

When G is O, we can rotate any vertex to any other vertex, any face to any other face, and any edge to any other edge. But we can't rotate between objects of different types — for example, we can't rotate a face to a vertex. So the poles form three orbits, based on whether they correspond to a face, vertex, or edge.

Now we can use this action to prove our classification.

Proof of Theorem 4.25. Consider the set S of pairs (g, p) such that g is not the identity, and p is a pole of g. The main idea is to count $|S|$ in two ways.

First, we'll count $|S|$ by looking at g. Every g other than the identity has exactly two poles, so we have

$$|S| = \sum_{g \neq e} 2 = 2(n-1). \tag{2}$$

On the other hand, we can also count by poles. For every p, the g for which p is a pole of g are exactly the elements of $\text{Stab}(p)$, other than the identity. So we have

$$|S| = \sum_p (|\text{Stab}(p)| - 1)$$

as well. Now we can group this sum by orbit — all stabilizers of poles in the same orbit have the same size, so for each orbit O_i there are n_i poles p each with $|\text{Stab}(p)| = r_i$, which means we can rewrite this as

$$|S| = \sum_{i=1}^k n_i(r_i - 1) = \sum_{i=1}^k \frac{n}{r_i}(r_i - 1). \tag{3}$$

Now setting our expressions for $|S|$ in (2) and (3) equal to each other and dividing by n, we get

$$\sum_{i=1}^k \left(1 - \frac{1}{r_i}\right) = 2 - \frac{2}{n}.$$

Now note that each stabilizer has size at least 2 — if p is a pole of g, then both g and the identity are in $\text{Stab}(p)$. So then $r_i \geq 2$ for all i, which means $1 - \frac{1}{r_i}$ is always in $[\frac{1}{2}, 1)$. Meanwhile, $2 - \frac{2}{n}$ is always in $[1, 2)$. This immediately implies we must have exactly 2 or 3 orbits (meaning k is 2 or 3)! This is already quite a strong constraint, and now we can split into cases.

Case 1 ($k = 2$). Then we have

$$1 - \frac{1}{r_1} + 1 - \frac{1}{r_2} = 2 - \frac{2}{n},$$

which means that

$$\frac{1}{r_1} + \frac{1}{r_2} = \frac{2}{n}.$$

But we must have $r_1, r_2 \leq n$, since the size of any stabilizer is at most the size of the entire group. So the only way for equality to hold here is if $r_1 = r_2 = n$. In that case, we have $n_1 = n_2 = 1$, so there's exactly two poles, and both are fixed by the entire group. This means all rotations are about the same axis. Then G is actually a finite subgroup of SO_2, which means it must be C_n.

Case 2 ($k = 3$). The equation we get is still quite constraining — we have

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

Without loss of generality, we can assume $r_1 \leq r_2 \leq r_3$. If $r_1 \geq 3$, then the left-hand side is at most $3 \cdot \frac{1}{3} = 1$, so then we must have $r_1 = 2$ (since it must be greater than 1).

Then if $r_2 \geq 4$, the left-hand side is again at most 1, so r_2 must be 2 or 3.

If $r_2 = 2$, then we get $r_2 = \frac{n}{2}$. Finally, if $r_2 = 3$, then we cannot have $r_3 \geq 6$ (or else the left-hand side would again be at most 1), so r_3 must be 3, 4, or 5.

So this gives a full classification of all possible stabilizer sizes: $(2, 2, r)$, $(2, 3, 3)$, $(2, 3, 4)$, and $(2, 3, 5)$. These correspond to n being $2r$, 12, 24, and 60, respectively.

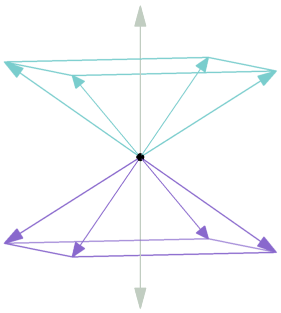
It's possible to show that the first case implies the group is D_r, the second implies the group is T (the rotational symmetries of a tetrahedron), the third implies the group is O (the rotational symmetries of an octahedron), and the fourth implies the group is I (the rotational symmetries of an icosahedron). Intuitively, it's unsurprising that we have three orbits in each case — for the symmetries of a regular polyhedron, we should have one orbit corresponding to the faces, one for the vertices, and one for the edges.

We won't work out the details in each of the cases — there's still some work to do, since it's possible that there could be different groups with the same orbit structure — but we'll show how to prove this for the case $(2, 3, 4)$, and the other cases can be handled similarly.

Suppose we have three orbits, with stabilizer sizes $(2, 3, 4)$. Then G has size 24, so we must have $n_3 = 6$. The six elements of this orbit are all unit vectors in three dimensions, so we can attempt to see what they look like.

Let one element of the orbit be p. Then $-p$ must be in the orbit as well — p and $-p$ have the same stabilizer, but here the stabilizers of elements in different orbits have different sizes. Then we must have $\text{Stab}(p) = \text{Stab}(-p) = C_4$, since the stabilizer has size 4 and consists of rotations about the axis through p.

Now let another element of the orbit be q. By the same reasoning, $-q$ is also in this orbit. But since our group contains C_4 (consisting of rotations about the axis through p), the rotations of q by multiples of 90° about this axis must all be in the orbit as well, and the same is true for $-q$.



Now if q were not perpendicular to p, this would give 10 distinct vectors in the orbit, but there can only be 6. So then q must be perpendicular to p (so that $-q$ is one of the rotations of q), and this accounts for all six vectors of the orbit.

But then G must fix this set of six vectors, so it must fix the octahedron whose vertices are the endpoints of these vectors. So then $G \leq O$; but both G and O have order 24, so we must have $G = O$. □

Remark 4.32. The hardest part of this proof was finding the idea to consider an action on the set of poles. Once we had the group action, we were then able to really strongly limit the possibilities for G by counting $|S|$ in two ways and messing around with the resulting equation.

§4.4 Conjugation

We'll now look at an action of G on *itself.* Of course, there's one obvious action — the ation $G \times G \to G$ sending $(g, x) \mapsto gx$. This action is transitive, and the stabilizer of any x consists of exactly the identity. So this is not very interesting.

But there's another, more interesting, way that G can act on itself — by *conjugation.* Here the action $G \times G \to G$ is defined by

$$(g, x) \mapsto gxg^{-1}.$$

It's possible to check that this satisfies the axioms for a group action.

§4.4.1 The Class Equation

Definition 4.33. The orbit of x under conjugation is called the *conjugacy class* of x and is denoted $C(x)$.

In other words, $C(x)$ is the set $\{gxg^{-1} \mid g \in G\}$.

Definition 4.34. The stabilizer of x under conjugation is called the *centralizer* of x and is denoted $Z(x)$.

In other words, $Z(x)$ is the set of $g \in G$ such that $gxg^{-1} = x$, or equivalently $gx = xg$ — so the centralizer of x consists of exactly the elements which commute with x.

All our theory about group actions still applies here — in particular, for all x we have

$$|G| = |C(x)| \cdot |Z(x)|.$$

We also know that the orbits partition the set, which here is G. This gives the *class equation*:

Proposition 4.35 (Class Equation)

If the conjugacy classes of G are $C_1, \ldots, C_n$, then $|G| = |C_1| + \cdots + |C_n|$, and each $|C_i|$ divides $|G|$.

Note also that $C(x)$ has size 1 (meaning it contains only x) if and only if $Z(x) = G$, meaning that x commutes with *every* element of G. Such elements have a name:

Definition 4.36. The *center* of G, denoted by Z, is the set of elements which commute with everything in G.

Using this, we can get constraints on $|C(x)|$ and $|Z(x)|$. For every x, both Z and $\langle x \rangle$ must be subgroups of $Z(x)$ (elements of Z commute with everything, and x commutes with itself). This can be quite powerful — the second constraint implies that $\text{ord}(x) \mid |Z(x)|$, and therefore

$$|C(x)| \,\Big|\, \frac{|G|}{\text{ord}(x)}.$$

Note also that any two elements in the same conjugacy class have the same order — since conjugation is an automorphism, we have $gx^kg^{-1} = (gxg^{-1})^k$, which means $x^k = e$ if and only if $(gxg^{-1})^k = e$.

We can often use these facts to constrain the class equation of a given group.

Example 4.37

What is the class equation of D_5?

Solution. We can write D_5 as $\{e, x, x^2, x^3, x^4, y, xy, x^2y, x^3y, x^4y\}$, where x is a rotation and y a reflection — so we have the relations $x^5 = y^2 = e$ and $yxy^{-1} = x^{-1} = x^4$.

First, we have $C(e) = \{e\}$, since the identity commutes with everything.

Next we look at orders — all elements in one conjugacy class must have the same order. The reflections (elements of the form x^ky) all have order 2, and the non-identity rotations (elements of the form x^k) all have order 5. So we cannot have a rotation and a reflection in the same conjugacy class.

First consider the reflection y. We have

$$\langle y \rangle \leq Z(y) \leq D_5,$$

so then $|Z(y)|$ must be a multiple of 2 and a divisor of 10. This means it must be either 2 or 10. But if $|Z(y)|$ were 10, then this would mean y commutes with the entire group, which is false. So then $|Z(y)| = 2$, which means $|C(y)| = 5$. So all the reflections are conjugate to each other — we have

$$C(y) = \{y, xy, x^2y, x^3y, x^4y\}.$$

Now we want to describe the conjugacy classes of the rotations — consider $C(x)$. We know $C(x)$ contains x, and it also contains x^4, since $x^4 = yxy^{-1}$. But $|C(x)|$ must divide 10, and it can be at most 4 (since we only have 4 elements left), so it can't contain any other elements. So we have

$$C(x) = \{x, x^4\} \text{ and } C(x^2) = \{x^2, x^3\}.$$

So the class equation is

$$10 = 1 + 5 + 2 + 2.$$

□

Remark 4.38. Note that the conjugacy classes of a group generally have different sizes — the behaviour of the partition into conjugacy classes is quite different from the behaviour of the partition into cosets.

§4.4.2 p-groups

One example of how conjugation can be useful is in analyzing p-groups.

Definition 4.39. Given a prime p, a group G is a *p-group* if $|G| = p^e$ for some $e \geq 0$.

Example 4.40

Any cyclic group C_{p^k} is a p-group, and so is any product of such cyclic groups.

Example 4.41

An example of a non-abelian p-group (with order p^3) is

$$\left\{\begin{bmatrix} 1 & * & * \\ 0 & * & * \\ 0 & 0 & 1 \end{bmatrix}\right\} \leq \mathrm{GL}_3(\mathbb{F}_p).$$

Theorem 4.42

Every p-group has nontrivial center.

In other words, this theorem states that $|Z| > 1$ — there is a non-identity element which commutes with everything. From our class equation for D_5 in Example 4.37, we can see that the center of D_5 is trivial — an element is in the center if and only if its conjugacy class has size 1, and the only conjugacy class of size 1 there is $\{e\}$. So the theorem states that this doesn't happen for p-groups — for example, the class equation of D_4 (which is a 2-group) is $8 = 1 + 1 + 2 + 2 + 2$, which means the center has size 2.

Proof. We use the class equation for G. We know that

$$p^e = |C_1| + \cdots + |C_k|,$$

where each $|C_i|$ divides p^e, and is therefore a power of p. But now grouping the conjugacy classes by size, we have

$$p^e = \underbrace{1 + 1 + \cdots + 1}_{|Z| \text{ times}} + (p + \cdots + p) + (p^2 + \cdots + p^2) + \cdots.$$

Now reducing mod p gives

$$0 \equiv |Z| \pmod{p}.$$

But $|Z| \geq 1$ (since Z necessarily contains e), so then $|Z| \geq p$, and Z is nontrivial. □

Example 4.43

The center of the p-group described in Example 4.41 consists exactly of matrices of the form

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

(this can be checked by explicit computation), which has order p.

Corollary 4.44

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

Proof. Since Z is a subgroup of G, its order must be p or p^2 (it must divie $|G|$, and it can't be 1 by the above theorem). If $|Z| = p^2$, then $Z = G$, and therefore G is abelian.

Now assume for contradiction that $|Z| = p$. then pick an element $x \in G$ which is *not* in Z, and look at the centralizer of x.

On one hand, $Z(x)$ must contain Z. But we can't have $Z(x) = Z$, since $Z(x)$ must also contain x (which is not in Z).

But $Z(x)$ is a subgroup of G, and $|Z(x)| > |Z| = p$, so then we must have $|Z(x)| = p^2$. This means x commutes with all elements of G, and therefore x must be in the center of G; this is a contradiction. □

The main point here is that p^2 is not very big, so there's not much room for things to happen. Note that Example 4.41 is a group of order p^3 which isn't abelian, so 2 is the largest exponent for which this statement holds.

In fact, it's possible to push this a bit further.

Proposition 4.45

If $|G| = p^2$, then G is isomorphic to either C_{p^2} or $C_p \times C_p$.

Recall that $C_p \times C_p$ consists of pairs of elements in C_p, with the operation performed componentwise.

Proof. We can consider the orders of elements in G, which must all divide p^2. First, if there exists some element a of order p^2, then G must be the cyclic group $\langle a \rangle$ (since $\langle a \rangle$ must be a subgroup of G, but they have the same size).

Otherwise, every element $a \neq e$ has order exactly p. We can now use the following general claim:

> **Claim** — If G is an abelian group such that every non-identity element has order exactly p, then we can think of G as a vector space over $\mathbb{F}_p$.

Proof. To turn G into a vector space, we can define addition using the addition operation in G (since we know that it's commutative). Meanwhile, we can define scalar multiplication by an element $\overline{n} \in \mathbb{F}_p$ by

$$\overline{n} \cdot g = \underbrace{g + g + \cdots + g}_{n \text{ times}}.$$

This is well-defined because g has order p, so two elements n and $n + p$ (which produce the same residue $\overline{n} \in \mathbb{F}^p$) also produce the same element $\overline{n}g$. With these two operations, G becomes a vector space. ■

Now in our situation, the dimension of the vector space must be 2 (since it must have exactly p^2 elements), so then the vector space is $\mathbb{F}_p^2$ and $G = C_p \times C_p$. □

§4.4.3 The Icosahedral Group

Another example of how we can use the class equation is in analyzing the icosahedral group I, the group of rotational symmetries of an icosahedron.

Earlier we mentioned that $|I| = 60$. To prove this, recall that all elements of I are rotations $\{\rho(\vec{u}, \theta)\}$, where the poles $\vec{u}$ correspond to the faces, edges, and vertices. Note that there is some redundancy — rotating by θ around $\vec{u}$ is the same as rotating by $-\theta$ around $-\vec{u}$. This pairs up the face rotations, edge rotations, and vertex rotations (note that this is different from the case of a tetrahedron, where if we have a pole corresponding to a vertex, its negative actually corresponds to a *face*).

We can now list all elements of I based on their axis of rotation. We then get the following rotations:

- The identity is in I.
- For each of the 20 faces, we can rotate by $2\pi/3$ or $4\pi/3$ about the corresponding axis, giving 2 nontrivial rotations.
- For each of the 30 edges, we can rotate by π around the axis to the center of the edge, giving 1 nontrivial rotation.
- For each of the 12 vertices, we can rotate around the axis to that vertex by a multiple of $2\pi/5$, giving 4 nontrivial rotations.

This overcounts every nontrivial rotation twice, so in total, $|I| = 1 + 20 + 15 + 24 = 60$.

> **Question 4.46.** What is the class equation for I?

We'll start by figuring out what conjugation does to I. Take some g and some $\rho(\vec{u}, \theta)$, and let

$$\rho = g\rho(\vec{u}, \theta)g^{-1}.$$

This must be a rotation by the same angle, around a different axis — we first rotate the icosahedron, then perform our original rotation, and then reverse the first rotation, which in total is the same angle of rotation (we can also see this by the fact that the angle of rotation is determined by the trace, which is preserved by conjugation). We can also describe what the new axis is — if the original rotation $\rho(\vec{u}, \theta)$ fixes p, then ρ fixes $g(p)$.

This implies that all face rotations by $2\pi/3$ are conjugate to each other. But rotation by $4\pi/3$ about an axis $\vec{u}$ is the same as rotation by $2\pi/3$ about the axis $-\vec{u}$. So then all 20 face rotations form one conjugacy class.

Similarly, there exist elements g mapping any edge to any other edge, so then the edge rotations are all conjugate to each other. So the 15 edge rotations make up another conjugacy class.

Finally, we consider the vertex rotations. We can again rotate any vertex to any other vertex. The rotations by $2\pi/5$ and $8\pi/5$ correspond to the same rotation angle, so these 12 vertex rotations are conjugate; similarly the rotations by $4\pi/5$ and $6\pi/5$ are also conjugate. So the class equation is

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

So I has 5 conjugacy classes, and its center has size 1 (since every element of the center corresponds to a conjugacy class of size 1 in the class equation).

We can now use this class equation to analyze the normal subgroups of I.

> **Definition 4.47.** A group G is *simple* if the only normal subgroups of G are $\{e\}$ and G itself.

Equivalently, G being simple means that any surjective homomorphism from G to another group is either an isomorphism or trivial (since the kernel of a homomorphism is always a normal subgroup).

Simple groups are important because in some sense, they're building blocks for all finite groups. If we have a group G that isn't simple, then we can write down a surjection $G \mapsto G'$, and analyze G in terms of the kernel and image of this surjection. We can analyze *those* by splitting them up again in this way, and so on — we keep decomposing the groups until they become simple (at which point we stop because we can't split them in a useful way).

> **Example 4.48**
>
> The cyclic group C_n is simple if and only if n is prime.

Theorem 4.49

The icosahedral group I is simple.

So I has a lot of subgroups, but it turns out that it doesn't have *any* interesting *normal* subgroups!

Proof. Suppose that N is a normal subgroup of I. Then $gNg^{-1} = N$ for all $g \in I$. So if an element x is in N, then its entire conjugacy class $C(x)$ is also in N; this means N is a union of conjugacy classes.

But we also know $|N|$ divides 60. Now consider the class equation

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

Then to build N, we must take 1 (since N must contain the identity) and some subset of the remaining terms. But we can check that the only ways to do so and end up with a factor of 60 are to take none or all of the remaining conjugacy classes, so we must have $|N|$ be 1 or 60. □

Remark 4.50. This proof is quite soft, in some sense. We don't have to grapple with the structure of N that much — we just look at its size in terms of the sizes of conjugacy classes.

Even when a group is *not* simple, it is sometimes possible to understand how to build normal subgroups by looking at the class equation. For example, we could use this to find the normal subgroups of D_5 (which is *not* simple, but doesn't have many normal subgroups).

The analysis of I has another interesting use.

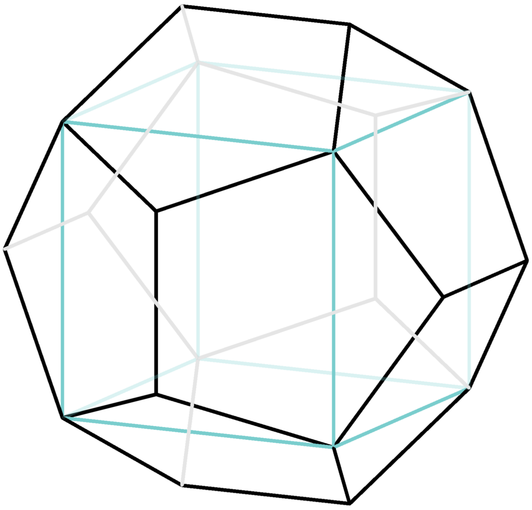
Theorem 4.51

The icosahedral group I is isomorhpic to the alternating group A_5.

Recall that A_5 is the subgroup of S_5 consisting of permutations with sign 1. Since S_5 has size $5! = 120$ and A_5 has index 2, then A_5 has size 60. So I and A_5 have the same size; but it turns out that they're actually the same group.

Proof. We want to describe I as a set of permutations of five objects, so we first want to find a group action of I on a set of size 5 — this will define a group homomorphism $I \to S_5$.

We can describe this action geometrically. Think of I as the group of rotational symmetries of a *dodecahedron* instead. Then inside this dodecahedron, we can produce five cubes, where the vertices of the cube are also vertices of the dodecahedron, and the edges of the cube are diagonals of the faces (which are all pentagons). There are five cubes because given any face, we have 5 choices for which diagonal of the face is used, and each uniquely determines the rest of the cube.



Let S be the set of these five cubes. Then rotating the dodecahedron must send one cube to another, so it gives a group homomorphism

$$f\colon I \to \mathrm{Perm}(S) = S_5.$$

This homomorphism is nontrivial — rotating around a face changes the diagonal we use from that face, which changes the corresponding cube. So its kernel cannot be I, which means its kernel is $\{e\}$, and therefore the homomorphism f is injective.

Now we want to show that the image of f is A_5. Consider the homomorphism $\varphi = \mathrm{sgn} \circ f$, which maps $I \to S_5 \to \{\pm 1\}$. Then $\ker(\varphi)$ must again be $\{e\}$ or I. But if $\ker(\varphi)$ were trivial, then φ would be injective — this is impossible because we can't have an injection from a set of size 60 to a set of size 2. So then $\ker(\varphi) = I$, which means that all elements of I are mapped to permutations with sign 1, and therefore

$$f(I) \subset \ker(\mathrm{sgn}) = A_5.$$

So then we can think of f as a homomorphism $I \to A_5$. But this homomorphism is injective, and I and A_5 have the same size, so it must be surjective as well. So f is an isomorphism between I and A_5. □

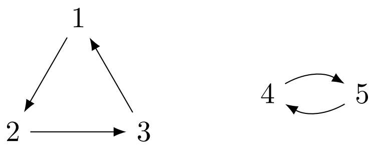
Corollary 4.52

The alternating group A_5 is simple.

In fact, A_n is simple for *all* integers $n \geq 5$. But the proof we saw here only works for $n = 5$; for larger n we really do need to get our hands dirty working with permutations and conjugacy classes.

§4.4.4 Conjugacy Classes of Permutations

We'll now consider the conjugacy classes in S_n and A_n. It'll often be useful to use *cycle notation*, where we write permutations as a product of disjoint cycles. For example, (123)(45) corresponds to the following permutation:



Elements which aren't mentioned in the cycle notation of a permutation are mapped to themselves — for example, if we considered $(123)(45) \in S_6$, then 6 would be mapped to itself.

Fact 4.53 — The sign of a permutation σ is $(-1)^k$, where k is the number of even-length cycles in σ.

Proof. By definition, if we can write σ as a product of n transpositions, then $\operatorname{sgn}(\sigma) = (-1)^n$. But we can write a m-cycle as

$$(123\cdots m) = (1m)\cdots(14)(13)(12),$$

so the sign of a m-cycle is $(-1)^{m-1}$. This means each even-length cycle in σ multiplies its sign by -1, and each odd-length cycle doesn't affect its sign. □

Equivalently, if σ has cycle lengths $k_1, \ldots, k_n$, then

$$\operatorname{sgn}(\sigma) = \prod_{i=1}^{n} (-1)^{k_i - 1}.$$

Question 4.54. What are the conjugacy classes in S_n?

It turns out that cycle notation is really good at describing conjugacy classes.

Example 4.55

Let $\sigma = (123)$, and take a permutation $p \in S_n$. What is the cycle notation of $\tau = p\sigma p^{-1}$?

Solution. Let $p(1) = i$, $p(2) = j$, and $p(3) = k$. Then we have

$$\tau(i) = p\sigma p^{-1}(p(1)) = p\sigma(1) = p(2) = j.$$

Similarly we have $\tau(j) = k$ and $\tau(k) = i$. We can use the same reasoning to check that τ fixes all other elements. So we have $\tau = (ijk)$ — the conjugate of our 3-cycle is another 3-cycle, with different elements. □

If we started off with a more complicated permutation for σ, the same thing would happen — for example, if $\sigma = (123)(47)\cdots$, then we would have

$$p\sigma p^{-1} = (p(1)p(2)p(3))(p(4)p(7))\cdots.$$

So the cycles in any conjugate of σ are the same as the cycles in σ, except with different numbers. To keep track of this more precisely, we can use the concept of *cycle type* — the cycle type of a permutation keeps track of the number of cycles of each length.

Proposition 4.56

Two permutations σ and τ are conjugate if and only if they have the same cycle type.

Proof. We've already seen one direction — if $\tau = p\sigma p^{-1}$, then the cycle notation of τ is obtained by writing down the cycle notation of σ and replacing each i with $p(i)$.

For the other direction, we can just match up corresponding cycles. For example, if $\sigma = (145)(23)$ and $\tau = (234)(15)$, then we can define p to be the permutation sending $1 \mapsto 2$, $4 \mapsto 3$, and so on. □

Example 4.57

In S_4, the conjugacy class of (1234) consists exactly of 4-cycles. To write down a 4-cycle, we can first write down some ordering of 1234 (in 24 ways); then this counts every cycle 4 times. So the conjugacy class has 6 elements.

There's another way to find the size of a conjugacy class — we can use the fact that $|C(x)| \cdot |Z(x)| = |G|$.

Example 4.58

To find the size of the conjugacy class of $x = (1234)$ in S_4, we can first find $Z(x)$. A permutation p is in $Z(x)$ if and only if $pxp^{-1} = x$, meaning that relabelling the cycle notation of x by replacing $i \mapsto p(i)$ doesn't change the permutation. We can relabel (1234) to any of (2341), (3412), and (4123). So then $|Z(x)| = 4$, and $|C(x)| = 24/4 = 6$.

Example 4.59

In S_{13}, what is the size of the conjugacy class of

$$x = (123)(456)(789\,10)(11)(12)(13)?$$

Solution. We'll start by finding $|Z(x)|$. We again want to find the number of ways to relabel the elements in this cycle notation which produce the same permutation. First, there's only one 4-cycle, so the relabelling of $(789\,10)$ must be the same cycle — then there's 4 ways to relabel it (since we have 4 choices of which element to write first).

For (11), (12), and (13), any reordering of these three elements will give the same permutation — for example, we could replace 11, 12, and 13 with 12, 11, and 13, and this would still correspond to the permutation fixing all of them. So this gives $3! = 6$ ways to relabel.

Finally, with the two 3-cycles (123) and (456), both of the above situations happen — for each cycle there's 3 different starting points, and we can also swap the two cycles in our relabelling .This gives $3 \cdot 3 \cdot 2! = 18$ ways.

So then we have $|Z(x)| = 4 \cdot 6 \cdot 18 = 432$, and therefore $|C(x)| = 13!\,/432$. □

We now know how to compute the sizes of conjugacy classes, which we can use to compute the class equation.

Example 4.60

To find the class equation of S_4, we can list all possible cycle types, calculate $|Z(x)|$ in the same way as above, and calculate $|C(x)|$ using the fact that $|C(x)| \cdot |Z(x)| = 24$:

Cycle Type	$\lvert Z(x)\rvert$	$\lvert C(x)\rvert$
4	4	6
$3+1$	3	8
$2+1+1$	$2 \cdot 2! = 4$	6
$2+2$	$2!\cdot 2 \cdot 2 = 8$	3
$1+1+1+1$	24	1

So S_4 has class equation

$$24 = 1 + 3 + 6 + 8 + 6.$$

We can also analyze the conjugacy classes of A_n. Since $A_n = \ker(\text{sgn})$ is a normal subgroup of S_n, it must be a union of conjugacy classes of S_n.

In the case of A_4, using Fact 4.53, the cycle types corresponding to even permutations are exactly $3+1$, $2+2$, and $1+1+1+1$, so the conjugacy classes of S_n which make up A_n give

$$|A_4| = 12 = 8 + 3 + 1.$$

However, note that this is *not* the class equation of A_4 — 8 doesn't divide 12, so we can't have a conjugacy class of size 8.

What went wrong is that two permutations σ and τ can be conjugate in S_n without being conjugate in A_n — if they're conjugate in S_n, then we know $\tau = p\sigma p^{-1}$ for some $p \in S_n$, but for them to be conjugate in A_n, we need $\tau = q\sigma q^{-1}$ for some $q \in A_n$. So if the only relabelling permutations p are odd, then τ and σ may no longer be conjugate in A_n.

However, the conjugacy classes in S_n and A_n are still closely related. Consider some $x \in A_n$. Then its conjugacy class in A_n must be a subset of its conjugacy class in S_n — we have

$$C_A(x) = \{y \in A \mid y = pxp^{-1} \text{ for some } p \in A_n\} \subset C_S(x) = \{y \in A_n \mid y = pxp^{-1} \text{ for some } p \in S_n\}.$$

Similarly, we can look at the stabilizers as well; then $Z_A(x) \leq Z_S(x)$, since any element of A_n which commutes with x is also an element of S_n which commutes with x.

But we also have

$$|C_A(x)| \cdot |Z_A(x)| = |A_n| = \frac{1}{2}|S_n| = \frac{1}{2}|C_S(x)| \cdot |Z_S(x)|.$$

Then since $Z_A(x)$ must *divide* $Z_S(x)$, we must either have $|C_A(x)| = |C_S(x)|$ and $|Z_A(x)| = \frac{1}{2}|Z_S(x)|$, or $|C_A(x)| = \frac{1}{2}|C_S(x)|$ and $|Z_A(x)| = |Z_S(x)|$. So each conjugacy class in S_n either remains the same or splits into two in A_n, and we just need to figure out which conjugacy classes split.

Example 4.61

In the case of A_4, we had $12 = 8 + 3 + 1$. The conjugacy class of size 8 must split because $8 \nmid 12$, while the conjugacy classes of sizes 3 and 1 cannot split (since their sizes are odd). So the class equation is

$$12 = 4 + 4 + 3 + 1.$$

Note that the second case $|Z_A(x)| = |Z_S(x)|$ occurs if and only if every permutation which commutes with x is even. Equivalently, the first case $|C_A(x)| = |C_S(x)|$ occurs if and only if there is an odd permutation commuting with x (since if the conjugay class doesn't split, then the stabilizer must shrink).

For example, in our conjugacy class of size 8 — which consists of 3-cycles — the conjugacy class does split, so all permutations commuting with (123) must be even. We could have checked this directly — the only permutations which commute with (123) are $\langle(123)\rangle = \{e, (123), (132)\}$, which are all even.

It's possible to perform a similar analysis for S_5 as well:

Example 4.62

In S_5 there are 7 conjugacy classes, giving the class equation

$$120 = 1 + 10 + 15 + 20 + 20 + 30 + 24.$$

The conjugacy classes corresponding to even permutations are 1, 15, 20, and 24, giving

$$60 = 1 + 15 + 20 + 24.$$

We know that 24 must split (since it doesn't divide 60), and 1 and 15 can't split. We can check that 20 doesn't split either, so the class equation of A_5 is

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

§4.5 The Sylow Theorems

Recall that if G is a finite group, and H is a subgroup of G, then $|H|$ divides $|G|$. In general, the converse is false — we can't necessarily find a subgroup of size d for every $d \mid |G|$. For example, A_4 has size 12, but it does not have a subgroup of order 6 (this can be checked using the class equation, as any such subgroup would have to be normal).

But surprisingly, this *is* true in general for certain values of d.

Theorem 4.63 (Sylow I)

Let G be a finite group with $|G| = n = p^e m$, where $p \nmid m$. Then there exists a subgroup $H \leq G$ such that $|H| = p^e$, called a *Sylow p-subgroup* of G.

Example 4.64

The group S_4 has order $24 = 8 \cdot 3$, so the first Sylow theorem implies there is a subgroup of order 8. One such subgroup is the subgroup $\langle (12), (34), (13), (24) \rangle$.

Example 4.65

The group D_5 has order $10 = 5 \cdot 2$, so the first Sylow theorem implies there is a subgroup of order 5 and a subgroup of order 2. One subgroup of order 5 is the subgroup C_5 generated by a rotation by $2\pi/5$, and one subgroup of order 2 is the subgroup D_1 generated by a reflection.

The power of this theorem is how general it is — amazingly, we can start off not knowing *anything* about the group, and even without knowing anything specific about the structure, we can know that a bunch of subgroups exist.

Corollary 4.66

For any prime p dividing $|G|$, there must exist an element of G with order p.

So for example, if we have a group of order 1, we *must* have an element of order 7 — we can't just have the identity and 13 elements of order 2.

Proof. From the first Sylow theorem, we know there exists a subgroup $H \leq G$ with $|H| = p^e$ for some e.

Then pick some y in H, so $\text{ord}(y) = p^f$ for some f. Now $x = y^{p^{f-1}}$ must have order p. □

Theorem 4.67 (Sylow II)

Any two Sylow p-subgroups are conjugate. More generally, for any Sylow p-subgroup $H \leq G$ and any p-group $K \leq G$, there exists $g \in G$ such that $gKg^{-1} \leq H$.

Example 4.68

In D_5, every reflection generates a Sylow 2-subgroup, and the second Sylow theorem points out that these subgroups are all conjugate to each other.

Note that the second statement implies the first (by taking K to be another Sylow p-subgroup).

Theorem 4.69 (Sylow III)

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

Example 4.70

In D_5, if we take $p = 2$, we have five Sylow 2-subgroups (one generated by each reflection); and 5 divides 5 and is 1 mod 2. Meanwhile, if we take $p = 5$, there's only one Sylow 5-subgroup, and 1 divides 2 and is 1 mod 5.

Before we prove these theorems, we'll see a few examples of how useful they are.

§4.5.1 Classifying Groups of Small Order

Since the Sylow theorems give us information about a group given only its size, we can use them to analyze what *all* groups of a given size can look like.

Example 4.71

There is only one group of size 15, up to isomorphism.

Proof. By the first Sylow theorem, we know G has a Sylow subgroup of order 5, and one of order 3.

First we can look at the subgroups of order 5. By the third Sylow theorem, the number of such subgroups must divide 3 and be 1 mod 5, which means it must equal 1. But if H is a subgroup of order 5, then all of its conjugates gHg^{-1} are also subgroups of order 5; so then these all must give the *same* subgroup, and H must be normal.

Similarly, we can look at the subgroups of order 3. The number of such subgroups must divide 5 and be 1 mod 3, so again there is exactly one subgroup K of order 3, which must be normal.

Note also that H and K cannot share any non-identity element, since the non-identity elements of H have order 5 and the non-identity elements of K have order 3.

Claim — The groups H and K commute — for any $h \in H$ and $k \in K$, we have $hk = kh$.

Proof. Since K is normal, we must have $hkh^{-1} \in K$, and therefore $hkh^{-1}k^{-1} \in K$ as well. Similarly since H is normal, $kh^{-1}k^{-1}$ must be in H, and therefore $hkh^{-1}k^{-1}$ must be in H as well. But then $hkh^{-1}k^{-1}$ must be in both K and H, so it must be the identity. ■

Claim — There is an isomorphism $H \times K \to G$.

Recall that $H \times K$ is the *product group* $H \times K = \{(h, k) \mid h \in H, k \in K\}$.

Proof. Consider the map $f \colon H \times K \to G$ sending $(h, k) \mapsto hk$. First we'll check that f is a homomorphism: for any (h, k) and (h', k') we have

$$f((h, k) \cdot (h', k')) = hk \cdot h'k' = hh' \cdot kk' = f(h, k)f(h', k'),$$

since H and K commute.

Now to show f is an isomorphism, note that $\ker(f)$ consists of elements (h, k) with $hk = 1$, which implies that $k = h^{-1}$ must be in H. But since K and H don't intersect except at 1, this requires that $h = k = 1$. So $\ker(f)$ is trivial, and therefore f is injective.

Finally, $H \times K$ and G both have order 15, so since f is injective, it must be a bijection. ■

So then any group of order 15 is isomorphic to $C_5 \times C_3$. □

Example 4.72

What are the possible groups of order 10 (up to isomorphism)?

Solution. We have $10 = 5 \cdot 2$. The number of Sylow 5-groups divides 2 and is 1 mod 5, so there must be exactly one Sylow 5-group K, and it must again be normal. Meanwhile, we know there exists some Sylow 2-group H, which may or may not be normal.

Both K and H must be cyclic, so we can write $K = \langle x \rangle$ where $\operatorname{ord}(x) = 5$, and $H = \langle y \rangle$ where $\operatorname{ord}(y) = 2$. Then we again have $K \cap H = \{1\}$, and since K is normal, we have $yxy^{-1} = x^r$ for some r.

Now we can try to write down elements of our group using x and y — our group must contain the elements

$$\{x^i y^j \mid 0 \le i \le 4, 0 \le j \le 1\}.$$

But this gives 10 distinct elements, so then G must consist of *exactly* these elements. So then we have

$$G = \langle x, y \mid x^5 = y^2 = e, yx = x^r y\rangle.$$

This completely determines the group, so it remains to figure out what values of r are possible. Every r gives at most one group, but some r may not work — for example, if $r = 2$ then we have

$$x = y^2 x = yyx = yx^2 y = x^4 y^2 = x^4,$$

so x has order dividing 3, contradiction. We can make the same argument in general: we have

$$x = y^2 x = yx^r y = x^{r^2} y^2 = x^{r^2}$$

by repeatedly applying $yx = x^r y$ to move the y's to the right, so then $x^{r^2-1} = 1$ and we must have $5 \mid r^2 - 1$, and therefore r must be 1 or 4.

So there's at most two possible groups. Both work — when $r = 1$ we get $xy = yx$, so K and H again commute and we get $C_5 \times C_2 = C_{10}$. Meanwhile, the case $r = 4$ gives the group D_5. So the only two groups of order 10 are C_{10} and D_5. □

Note that the situation for 10 was somewhat more subtle than for 15 because the Sylow theorems only guaranteed that *one* of the subgroups was normal.

In general, we can use the same argument to analyze groups with order pq for distinct primes $p < q$. If $q \not\equiv 1 \pmod{p}$, then similarly to Example 4.71 (where we had order $3 \cdot 5 = 15$) the only group up to isomorphism is $C_p \times C_q = C_{pq}$. Meanwhile, if $q \equiv 1 \pmod{p}$, then similarly to Example 4.72 (where we had order $2 \cdot 5 = 10$) there will be two possible groups — C_{pq} and some non-abelian group. The proof is the exact same — we look at the Sylow p-groups and q-groups.

§4.5.2 Classifying Abelian Groups

Question 4.73. What can we say about finite abelian groups?

It turns out that we can perform an analysis very similar to the one we did in Example 4.71 for general groups of order 15.

Suppose G is a finite abelian group with order $p_1^{e_1} \cdots p_r^{e_r}$. Then for each prime, we can take a Sylow subgroup H_i with order $p_i^{e_i}$. The choice of each H_i must be unique — any two Sylow p_i-subgroups must be conjugate, but conjugation doesn't have any effect since the group is abelian.

Now we can consider the product group $H_1 \times \cdots \times H_r$, which is again an abelian group. Define the map $f\colon H_1 \times \cdots \times H_r \to G$ by $(x_1, \dots, x_r) \mapsto x_1 + \cdots + x_r$. (Here we use addition instead of multiplication to denote the group operation because the group is abelian.)

Claim — f is an isomorphism.

Proof. First, f is a homomorphism because G is abelian, so all elements commute. But now note that $H_i \leq \mathrm{im}(f) \leq G$ for all i, which means $p_i^{e_i}$ must divide $|\mathrm{im}(f)|$ for all i. But then their product $\prod p_i^{e_i} = |G|$ must divide $|\mathrm{im}(f)|$ as well, which means $\mathrm{im}(f) = G$ and f is surjective.

On the other hand, $H_1 \times \cdots \times H_r$ and G have the same size. So since f is surjective, it must be injective as well, and therefore f is an isomorphism. □

This gives the following result:

Proposition 4.74

Any finite abelian group is isomorphic to a product of abelian groups with prime power order.

So then in order to understand finite abelian groups, it's enough to understand finite abelian groups of order p^k for primes p. In fact, this is fully understood, and we'll see a full classification in **18.702**.

§4.5.3 Proofs of Sylow Theorems

Now we'll prove the Sylow theorems. The main idea in all of the proofs will be to find a useful action of G on some set, and exploit this action to get information about G. We've been doing this for the past few weeks, but the main difference is that here we don't know anything about the group to start with.

We'll start with the first Sylow theorem.

Proof of Theorem 4.63. First, we need a set S that G acts on — take S to be the set of subsets of G with size p^e, so $|S| = \binom{n}{p^e}$. Then G acts on S by left translation — an element g maps a subset U to the subset

$$gU = \{gu \mid u \in U\}.$$

Our goal is to find a subgroup of a certain size. We can obtain subgroups from a group action by looking at stabilizers, and we can analyze the sizes of stabilizers by analyzing the sizes of the corresponding orbits.

> **Fact** — $\binom{n}{p^e}$ is not a multiple of p.

This is possible to prove just by writing out the explicit formula and counting the powers of p in the numerator and denominator. In fact, a stronger statement is true — we always have $\binom{n}{p^e} \equiv m \pmod{p}$ — but we won't need this here.

Lemma 4.75

If U is a subset of G and $H \le G$ stabilizes U, then $|H|$ divides $|U|$.

Proof. Since H stabilizes U, then by definition, for every $h \in H$ and $u \in U$ we must have $hu \in U$ as well. But this means for each $u \in U$, the coset Hu is contained in U, so then we can partition U into right cosets of H. Each coset has size $|H|$, so then $|H|$ must divide $|U|$. (The reason it's important here that H stabilizes U is because otherwise, the cosets of H may not be contained in U.) ■

Now we'll find a Sylow p-subgroup by looking at stabilizers. Consider the partition of S into orbits, with

$$|S| = |O_1| + \cdots + |O_r|.$$

Then since $|S|$ is not divisible by p, some orbit must also have size not divisible by p. Let this orbit be O, and let U be some element of the orbit; then we have

$$|G| = p^e m = |O| \cdot |\mathrm{Stab}(U)|.$$

But p doesn't divide $|O|$, so then p^e must divide $|\mathrm{Stab}(U)|$. On the other hand, by the above lemma, $|\mathrm{Stab}(U)|$ must divide $|U| = p^e$. So then we must have $|\mathrm{Stab}(U)| = p^e$, which means $|\mathrm{Stab}(U)|$ is a Sylow p-subgroup. □

> **Remark 4.76.** This is a very clever proof. The most important leap is the first one — picking the set S on which the group acts. As we've seen many times so far, group actions can be really useful, and finding a good set for a group action can take some amount of trial and error.

Now we'll prove the second Sylow theorem — more precisely, we'll show that given any Sylow p-subgroup H, any p-group $K \le G$ must be conjugate to a subgroup of H.

Proof of Theorem 4.67. Fix the Sylow p-subgroup H and the p-group $K \le G$, and let $|K| = p^f$.

Let X be the set of left cosets of H in G, so $|X| = m$. Now consider the action of K on X given by left translation — an element k maps $aH \mapsto kaH$. We again have a set and a group acting on it, so we can decompose the set into orbits, giving

$$m = |X| = |O_1| + \cdots + |O_r|.$$

Then each $|O_i|$ must divide $|K| = p^f$, but p doesn't divide m, so some orbit O must have size 1. This means some coset $aH \in X$ is fixed by all elements of K.

But then $kaH = aH$ for all $k \in K$. This means $a^{-1}kaH = H$ for each k, and therefore $a^{-1}ka \in H$ for all $k \in K$. So then $a^{-1}Ka$ is a subgroup of H, as desired. □

> **Remark 4.77.** Again, we see the common theme that when we have a group action, it's really useful to look at the orbit decomposition.

Finally, we'll prove the third Sylow theorem — this is the sneakiest proof.

Proof of Theorem 4.69. We again consider a group acting on a set. This time, our set Y will be the set of Sylow p-subgroups of G (so we're interested in $|Y|$).

To prove the first part of the theorem, consider the action of G on Y by conjugation — an element g maps $H \mapsto gHg^{-1}$ (which is another Sylow p-subgroup). Then by the second Sylow theorem, this action only has one orbit. So for any Sylow p-subgroup H, we have

$$|G| = |O(H)| \cdot |\text{Stab}(H)| = |Y| \cdot |\text{Stab}(H)| .$$

This immediately tells us that $|Y|$ must divide n, and we can gain more information from analyzing $|\text{Stab}(H)|$.

By definition, $\text{Stab}(H)$ is the set of g for which $gHg^{-1} = H$. But it's clear that $H \leq \text{Stab}(H)$, since for any $h \in H$ we have $hHh^{-1} = H$. So then $|\text{Stab}(H)|$ is divisible by $|H| = p^e$. Then

$$p^e m = |Y| \cdot p^e \ell$$

for some ℓ, which means $|Y|$ must divide m. This proves the first part of the theorem.

Now for the second part, we take the same set Y, but instead of considering the action of G, we instead consider the action of H, where H is any Sylow p-subgroup. (The action is still given by conjugation.)

Lemma 4.78

A Sylow p-subgroup $H' \in Y$ is fixed under conjugation by H if and only if $H = H'$.

In other words, we have $hH'h^{-1} = H'$ for all $h \in H$ if and only if $H' = H$.

Proof. It's clear that H is fixed under conjugation by any of its elements, so it suffices to show the other direction — that if H' is fixed by conjugation by all elements of H, then we must have $H' = H$.

Consider the group action of G on Y again, and consider the set $\text{Stab}_G(H')$, which consists of $g \in G$ for which $gH'g^{-1} = H'$. This is also denoted $N(H')$ and called the *normalizer* of H'.

Then we must have $H' \leq N(H')$. Meanwhile since H' is fixed by H, we must have $H \leq N(H')$ as well.

Then since $N(H')$ is a subgroup of G, its order must be a multiple of $|H| = p^e$ and must divide $|G| = p^e m$, so the power of p dividing $N(H')$ is exactly p^e. But this means H and H' are both Sylow p-subgroups of $N(H')$ as well! Then we can use the second Sylow theorem — there must exist some element $n \in N(H')$ such that $nH'n^{-1} = H$. But by the definition of $N(H')$, we must have $nH'n^{-1} = H'$ for any $n \in N(H')$, so then we must have $H = H'$. ■

So then our group action by H on Y has exactly one fixed point, which is H. We can again consider the decomposition into orbits, which gives

$$|Y| = |O_1| + \cdots + |O_r| .$$

Each orbit has size dividing $|H| = p^e$, and exactly one has size 1, so all the other orbits have size divisible by p, and therefore $|Y| \equiv 1 \pmod p$. □

> **Remark 4.79.** This proof was quite sneaky because we used the previously shown Sylow theorems on a *different* group in order to prove the third Sylow theorem on G.

§5 Bilinear and Hermitian Forms

§5.1 Bilinear Forms

Let V be a vector space over $\mathbb{R}$.

Definition 5.1. A *bilinear form* is a function $V \times V \to \mathbb{R}$, denoted by $(v, w) \mapsto \langle v, w\rangle$, such that:

(1) $\langle v, cw\rangle = c\langle v, w\rangle$ for all scalars c,

(2) $\langle v, w_1 + w_2\rangle = \langle v, w_1\rangle + \langle v, w_2\rangle$,

(3) $\langle cv, w\rangle = c\langle v, w\rangle$ for all scalars c,

(4) $\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle$.

The properties (1) and (2) mean the form is linear in the second variable, and the properties (3) and (4) mean it's linear in the first variable — this is why the form is called *bilinear*.

Example 5.2

The function $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ defined as

$$\langle (x_1, x_2, x_3)^\intercal, (y_1, y_2, y_3)^\intercal\rangle = x_1y_1 + 2x_1y_2 + 3x_2y_1 + 4x_2y_3 + 5x_3y_1$$

is a bilinear form.

Intuitively, all bilinear forms should look something like this example — all terms should be of the form cx_iy_j, and we shouldn't have any constant terms or higher order terms.

Definition 5.3. A bilinear form is *symmetric* if $\langle v, w\rangle = \langle w, v\rangle$ for all vectors v and w.

Example 5.4

The form given in Example 5.2 is *not* symmetric, while the form

$$\langle (x_1, x_2, x_3)^\intercal, (y_1, y_2, y_3)^\intercal\rangle = x_1y_1 + 2x_2y_1 + 2x_1y_2 + 3x_2y_2$$

is symmetric.

Example 5.5

The dot product is a symmetric bilinear form.

§5.1.1 Bilinear Forms in Matrices

Similarly to in the case of linear transformations, we can explicitly describe bilinear forms using matrices.

Suppose our vector space is $\mathbb{R}^n$. Then the dot product is a symmetric bilinear form. More generally, given any $n \times n$ matrix A, we can define a bilinear form

$$\langle \vec{x}, \vec{y}\rangle = \vec{x}^\intercal A\vec{y}$$

(here $\vec{x}$ and $\vec{y}$ are column vectors, and the output is a real number). This satisfies the axioms for a bilinear form because of properties of matrix multiplication.

Definition 5.6. A $n \times n$ matrix is *symmetric* if it equals its transpose.

Given a general matrix A, for any $\vec{x}$ and $\vec{y}$ we have

$$\langle \vec{y}, \vec{x} \rangle = \vec{y}^\intercal A \vec{x} = (\vec{y}^\intercal A \vec{x})^\intercal = \vec{x}^\intercal A^\intercal \vec{y}.$$

So then if $A^\intercal = A$, we have $\langle \vec{y}, \vec{x} \rangle = \langle \vec{x}, \vec{y} \rangle$, and the form is symmetric. Similarly to the case of linear transformations, the converses of both statements are true:

Proposition 5.7

Every bilinear form on $\mathbb{R}^n$ corresponds to a matrix — given a form $\langle -, - \rangle$, there is a unique matrix A such that $\langle \vec{x}, \vec{y} \rangle = \vec{x}^\intercal A \vec{y}$. Furthermore, $\langle -, - \rangle$ is symmetric if and only if A is.

Example 5.8

The dot product corresponds to the identity matrix.

Example 5.9

The bilinear form in Example 5.2 can be written as

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^\intercal \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 4 \\ 5 & 0 & 0 \end{bmatrix} \vec{y},$$

and the bilinear form in Example 5.4 can be written as

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^\intercal \begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \vec{y}.$$

We can see that the entries of the matrix come from the coefficients of the form.

Proof of Proposition 5.7. Let $\vec{e_1}, \dots, \vec{e_n}$ be the standard basis of $\mathbb{R}^n$ (so $\vec{e_i}$ has a 1 in its ith position and 0's everywhere else). Then let $a_{ij} = \langle \vec{e_i}, \vec{e_j} \rangle$ and create a matrix $A = (a_{ij})$. Now for any $\vec{x} = \sum x_i \vec{e_i}$ and $\vec{y} = \sum y_i \vec{e_i}$, by using bilinearity we have

$$\langle x, y \rangle = \left\langle \sum_i x_i \vec{e_i}, \sum_j y_j \vec{e_j} \right\rangle = \sum_i \sum_j x_i \langle \vec{e_i}, \vec{e_j} \rangle y_j = \sum_i \sum_j x_i a_{ij} y_j.$$

But this is exactly $\vec{x}^\intercal A \vec{y}$, so every bilinear form is of the form $\vec{x}^\intercal A \vec{y}$.

Then $\langle -, - \rangle$ is symmetric if and only if $\langle \vec{e_i}, \vec{e_j} \rangle = \langle \vec{e_j}, \vec{e_i} \rangle$ for all i and j, or equivalently if $a_{ij} = a_{ji}$ for all i and j. This occurs exactly when A is symmetric. □

§5.1.2 Choosing a Basis

We've now seen how to describe bilinear forms in $\mathbb{R}^n$, and as usual we can use this to describe bilinear forms in any vector space V by picking a basis $\{v_1, \dots, v_n\}$ of V.

Once we fix a basis, Proposition 5.7 is true for a general vector space as well — every bilinear form corresponds to a matrix, and the bilinear form is symmetric if and only if the matrix is. More explicitly, we take the matrix $A = (a_{ij})$ where $a_{ij} = \langle v_i, v_j \rangle$, and then for two vectors v and w with coordinates $\vec{x}$ and $\vec{y}$ in our basis, we have $\langle v, w \rangle = \vec{x}^\intercal A \vec{y}$.

In some ways, this may remind us of the situation when we studied linear operators — we started with a linear operator $T\colon V \to V$, and by picking a basis for V, we turned the operator into a $n \times n$ matrix. Here we start with a bilinear form, and by picking a basis we can also turn the *form* into a $n \times n$ matrix. But these correspondences are quite different in some ways.

Question 5.10. What happens to the matrix if we change basis?

Suppose we have two bases $B\colon \mathbb{R}^n \to V$ and $B'\colon \mathbb{R}^n \to V$, so that $B' = BP$ for an invertible matrix P.

$$\begin{tikzcd}
 & v & \\
 & V & \\
\mathbb{R}^n \arrow[ur, "B"] & & \mathbb{R}^n \arrow[ul, "B'"'] \arrow[ll, "P"] \\
P\vec{x} & & \vec{x}
\end{tikzcd}$$

Then if a pair (v, w) corresponds to coordinates $(\vec{x}, \vec{y})$ in B', it corresponds to $(P\vec{x}, P\vec{y})$ in B. So if our form corresponds to A' in the basis B' and A in the basis B, we have

$$\vec{x}^\intercal A' \vec{y} = (P\vec{x})^\intercal A (P\vec{y}) = \vec{x}^\intercal P^\intercal A P \vec{y}.$$

So then when we change basis, the new matrix is $A' = P^\intercal A P$.

Remark 5.11. Note that this relation is different from the one we get for linear operators, where we have $A' = P^{-1}AP$.

In particular, we can check explicitly that if A is symmetric, then so is A'. This is unsurprising, since symmetry is a property of the bilinear form itself, and shouldn't depend on which basis is used.

Now that we know how to change basis, we can ask a question similar to the one we asked for linear operators:

Question 5.12. Given a vector space V with bilinear form $\langle -, - \rangle$, how nice can we make the corresponding matrix A by choosing a basis?

We'll return to this question later; it turns out that for *symmetric* bilinear forms, the answer is very nice.

§5.2 Hermitian Forms

So far, we've worked over the field $\mathbb{R}$, but we can also work over the field $\mathbb{C}$.

The definitions in the previous section still work over any field. But there's a special property that they have in $\mathbb{R}$ — the dot product has the property that $\vec{x} \cdot \vec{x} \geq 0$ for all vectors $\vec{x}$.

Definition 5.13. A symmetric bilinear form $\langle -, - \rangle$ is *positive definite* if $\langle v, v \rangle > 0$ for all nonzero vectors v.

The dot product in $\mathbb{R}^n$ lets us talk about the lengths of vectors, so we'd like some version of this property in a form over the complex numbers as well. It turns out that if we're willing to loosen the bilinearity restriction a bit, then there's a way of doing this.

First we can extend the dot product in a way that captures our notion of distance — for a single complex number, we have $|z|^2 = z \cdot \overline{z}$. So in general, we can use *complex conjugation* to define our form:

Definition 5.14. The *standard Hermitian form* on $\mathbb{C}^n$ is the map $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ given by

$$\langle \vec{x}, \vec{y} \rangle = \overline{x_1} \cdot y_1 + \overline{x_2} \cdot y_2 + \cdots + \overline{x_n} \cdot y_n.$$

This is similar to the dot product, but we conjugate all entries of the first vector. The standard Hermitian form for $\mathbb{C}$ has the same property as the dot product for $\mathbb{R}$ — when $\vec{y} = \vec{x}$, we get

$$\langle \vec{x}, \vec{x} \rangle = \overline{x_1} \cdot x_1 + \cdots + \overline{x_n} \cdot x_n = |x_1|^2 + \cdots + |x_n|^2,$$

which is a positive real number whenever $\vec{x}$ is nonzero. So then we'll use the standard Hermitian form as our central example for a form on $\mathbb{C}$-vector spaces, similarly to how the dot product was our central example for a form on $\mathbb{R}$-vector spaces.

To describe this construction a bit more efficiently, we'll use the following definition:

Definition 5.15. For a matrix M over $\mathbb{C}$, the *adjoint* of M, denoted M^*, is the matrix $\overline{M^\intercal}$.

The adjoint has similar properties to the transpose — in particular, we have $(AB)^* = B^*A^*$ for any matrices A and B. Using this notation, the standard Hermitian form can be described as

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^* \vec{y}.$$

Notice that in the standard Hermitian form, we have

$$\langle \alpha\vec{x}, \vec{y} \rangle = \overline{\alpha}\langle \vec{x}, \vec{y} \rangle,$$

instead of $\alpha\langle \vec{x}, \vec{y} \rangle$. So because of the complex conjugation, the standard Hermitian form isn't exactly linear in the first entry.

With this in mind, let's now define a general form for a $\mathbb{C}$-vector space.

Definition 5.16. For a vector space V over $\mathbb{C}$, a *Hermitian form* is a function $V \times V \to \mathbb{C}$ denoted by $(v, w) \mapsto \langle v, w \rangle$, such that:

(1) $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$,

(2) $\langle v, \alpha w \rangle = \alpha \langle v, w \rangle$ for all scalars α,

(3) $\langle w, v \rangle = \overline{\langle v, w \rangle}$.

So Hermitian forms over $\mathbb{C}$ are somewhat similar to symmetric bilinear forms over $\mathbb{R}$, except that we have complex conjugation thrown in.

Note that if $\langle -, - \rangle$ is any Hermitian form, then

$$\langle \alpha v, w \rangle = \overline{\langle w, \alpha v \rangle} = \overline{\alpha \langle w, v \rangle} = \overline{\alpha}\langle v, w \rangle,$$

which is the same property we observed earlier for the standard Hermitian form. In particular, we have the following important property:

Fact 5.17 — For any Hermitian form $\langle -, - \rangle$, we have $\langle v, v \rangle \in \mathbb{R}$ for all vectors v.

Proof. We have $\langle v, v \rangle = \overline{\langle v, v \rangle}$, and a complex number is self-conjugate if and only if it is real. □

§5.2.1 Hermitian Matrices

Question 5.18. What does the matrix for a Hermitian form look like?

As in the case of bilinear forms, given a Hermitian form $\langle -, - \rangle$ on V, we can choose a basis $\{v_1, \dots, v_n\}$ of V and set A to be the matrix consisting of the entries $a_{ij} = \langle v_i, v_j \rangle$ for all i and j. Then for any $v = x_1v_1 + \cdots + x_nv_n$ and $w = y_1v_1 + \cdots + y_nv_n$, we must have

$$\langle v, w \rangle = \vec{x}^* A \vec{y}.$$

So we can again describe the Hermitian form by a matrix (determined by the form's behaviour on the basis vectors), but here we use the adjoint rather than the transpose. We also have the condition that

$$a_{ij} = \langle v_i, v_j \rangle = \overline{\langle v_j, v_i \rangle} = \overline{a_{ji}}$$

for each i and j, which means $A^* = A$.

Definition 5.19. A matrix A is called a *Hermitian matrix* if $A^* = A$.

So then Hermitian forms on $\mathbb{C}^n$ are the same as Hermitian matrices.

When we looked at bilinear forms over $\mathbb{R}$, we saw that changing our basis changes the corresponding matrix from A to $P^\intercal AP$ for an invertible matrix P. Unsurprisingly, for Hermitian forms, changing basis changes the matrix to P^*AP.

Example 5.20

One example of a 2×2 Hermitian matrix is

$$A = \begin{bmatrix} 5 & 2+2i \\ 2-2i & 3 \end{bmatrix}.$$

This corresponds to the Hermitian form

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^* A \vec{y} = 5\overline{x_1} \cdot y_1 + 3\overline{x_2} \cdot y_2 + (2+2i)\overline{x_1} \cdot y_2 + (2-2i)\overline{x_2} \cdot y_1.$$

Note that $\langle \vec{x}, \vec{x} \rangle$ is always real, as

$$\langle \vec{x}, \vec{x} \rangle = 5\left|x_1\right|^2 + 5\left|x_2\right|^2 + 2\operatorname{Re}((2+2i)\overline{x_1} \cdot y_2).$$

Note that for any Hermitian matrix, the entries on the diagonal must all be real (as they must equal their own conjugates).

It turns out that Hermitian matrices have nice properties. We'll see more of these properties later, but here's one:

Proposition 5.21

If A is Hermitian, then all its eigenvalues are real.

Proof. Suppose $\lambda \in \mathbb{C}$ is an eigenvalue of A, so we have $A\vec{v} = \lambda\vec{v}$ for some $\vec{v} \in \mathbb{C}^n$. Then we have

$$\vec{v}^* A \vec{v} = \vec{v}^* \lambda \vec{v} = \lambda \vec{v^*} \vec{v}.$$

But $\vec{v}^* A \vec{v}$ is real since A is Hermitian, and $\vec{v}^* \vec{v}$ is also real and nonzero. So λ must be real — in fact, it's the ratio between the pairing $\langle \vec{v}, \vec{v} \rangle$ given by A and the pairing given by the standard Hermitian form. □

Even if all entries of A are real, this is a nontrivial fact — in general, a matrix with real entries can still have complex eigenvalues. But this guarantees that for *symmetric* real matrices, all their eigenvalues are necessarily real!

§5.3 Orthogonality

We'll now study symmetric bilinear forms in $\mathbb{R}$ and Hermitian forms in $\mathbb{C}$ in parallel, since the theory in the two cases is quite similar.

Earlier, when working with the vector space $\mathbb{R}^n$, we defined a matrix to be *orthogonal* if it preserves the dot product — meaning that $M\vec{x} \cdot M\vec{y} = \vec{x} \cdot \vec{y}$ for all vectors $\vec{x}$ and $\vec{y}$. We saw that this is equivalent to $M^\intercal M = I$, or to the column vectors $\vec{v_i}$ of M being orthonormal (meaning that $\vec{v_i} \cdot \vec{v_j}$ is 1 if $i = j$ and 0 otherwise). We can define a similar notion in the case of $\mathbb{C}^n$:

Definition 5.22. A matrix M with entries in $\mathbb{C}$ is *unitary* if for all vectors $\vec{x}, \vec{y} \in \mathbb{C}^n$ we have $\langle M\vec{x}, M\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle$, where $\langle -, - \rangle$ denotes the standard Hermitian form on $\mathbb{C}^n$.

Similarly to the case of orthogonal matrices, a matrix M is unitary if and only if $M^*M = I$, or equivalently if its column vectors $\vec{v_i}$ are orthonormal, again meaning that $\langle \vec{v_i}, \vec{v_j} \rangle$ is 1 if $i = j$ and 0 otherwise.

§5.3.1 Orthogonal Complements

In $\mathbb{R}^n$, the dot product gives us a way of describing when two vectors are perpendicular. In a general vector space, if we have a pairing "similar to" the dot product — a symmetric bilinear form for $\mathbb{R}$, or a Hermitian form for $\mathbb{C}$ — then we can use that form to define perpendicularity in the same way. So we'll now assume that V is either a real vector space with a symmetric bilinear form, or a complex vector space with a Hermitian form; we'll denote this form by $\langle -, - \rangle$.

Definition 5.23. Two vectors v and w are *orthogonal*, denoted as $v \perp w$, if $\langle v, w \rangle = 0$.

We can also describe when a vector is perpendicular to a *subspace* — given a subspace $W \subset V$, we say $v \perp W$ if $\langle v, w \rangle = 0$ for all $w \in W$.

The dot product in $\mathbb{R}^n$ captures our geometric notion of perpendicularity. However, this won't necessarily be true in general. In particular, a vector may be perpendicular to itself!

Example 5.24

Consider the bilinear form on $\mathbb{R}^4$ defined by

$$A = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

which comes up in special relativity. Then $(1, 0, 0, 1)^\intercal$ is orthogonal to itself.

Definition 5.25. Given a subspace $W \subset V$, we define its *orthogonal complement* $W^\perp$ as the subspace $\{v \in V \mid v \perp W\}$.

Question 5.26. When can we split V as a direct sum of W and $W^\perp$?

In $\mathbb{R}^3$, we always have $\mathbb{R}^3 = W \oplus W^\perp$ — for instance, if W is a plane, then $W^\perp$ is the line perpendicular to all lines in the plane. We'd *like* to make a similar statement for a general vector space and symmetric bilinear form. But this isn't true — we may have a vector v which is orthogonal to *all* of V. So we'd like to impose a condition on our form to avoid such behaviour, and for that we'll use the following definition:

Definition 5.27. The *null space* of $\langle -, - \rangle$ is the space of vectors v for which $v^\perp = V$.

In other words, the null space is the subspace of V consisting of vectors orthogonal to all of V. We'll denote the null space by N.

Example 5.28

In the form given by $A = I$ (the dot product or standard Hermitian form), we have $N = \{0\}$. Meanwhile in the form given by $A = 0$, we have $N = V$.

Definition 5.29. If $N = \{0\}$, then we say that $(V, \langle -, - \rangle)$ is *nondegenerate.*

So the form corresponding to I is nondegenerate, while the form corresponding to 0 is degenerate (as we would expect). More generally, we can describe the null space explicitly using the matrix A — we have $v \in N$ if and only if $w^* Av = 0$ for all $w \in V$. This occurs if and only if $Av = 0$, meaning that $v \in \ker(A)$. In particular, the pairing is nondegenerate if and only if A is invertible.

But V being nondegenerate doesn't exactly guarantee that the pairing is "well-behaved" — the pairing in Example 5.24 is nondegenerate, but it still has the weird behaviour that some vectors are perpendicular to themselves. This is because even if $\langle -, - \rangle$ is nondegenerate on V, its *restriction* to a subspace $W \subset V$ may be degenerate.

Example 5.30

The pairing in Example 5.24 is nondegenerate as a pairing on $\mathbb{R}^4$. But it becomes degenerate if we restrict it to $\text{Span}((1, 0, 0, 1)^\intercal)$, since any two vectors in this span are orthogonal.

By definition, the restriction of $\langle -, - \rangle$ to W is nondegenerate if and only if for all $w \in W$, there exists another vector $w' \in W$ such that $\langle w, w' \rangle = 0$. This is equivalent to stating that $W \cap W^\perp = \{0\}$ — if we had a vector w in both W and $W^\perp$, then by the definition of $W^\perp$, w would have to be orthogonal to all of W.

Now we're ready to answer the question on when we can use $\langle -, - \rangle$ to split V as a direct sum:

Theorem 5.31

If the restriction of $\langle -, - \rangle$ to W is nondegenerate, then $V = W \oplus W^\perp$.

Recall that the statement $V = W \oplus U$ means every $v \in V$ can be written uniquely as a sum $w + u$ where $w \in W$ and $u \in U$.

It's clear that the nondegeneracy condition is necessary — if $\langle -, - \rangle$ were degenerate, then we would have $W \cap W^\perp \neq \{0\}$, and therefore we could not have $V = W \oplus W^\perp$.

Proof. First, the condition that $\langle -, - \rangle$ restricted to W is nondegenerate means that $W \cap W^\perp = \{0\}$, so it suffices to show that $V = W + W^\perp$ — or equivalently, that every vector $v \in V$ can be written as the sum of a vector in W and one in $W^\perp$.

We'll work over $\mathbb{C}$, but the same argument works over $\mathbb{R}$ as well.

Pick a basis $\{w_1, \dots, w_k\}$ for W, and define a map $\varphi\colon V \to \mathbb{C}^k$ sending

$$v \mapsto (\langle w_1, v\rangle, \dots, \langle w_k, v\rangle).$$

Then ϕ is linear by the properties of Hermitian forms. Meanwhile, $\ker(\varphi)$ is the set of $v \in V$ for which $\langle w_i, v\rangle = 0$ for all the basis vectors w_i, but since the vectors w_i span W, this is true if and only if $\langle w, v\rangle = 0$ for *all* vectors $w \in W$. So then $\ker(\varphi) = W^\perp$.

Now we can use the dimension formula — we have

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

But we know $\ker(\varphi) = W^\perp$, and $\dim\operatorname{im}(\varphi) \le \dim W$ (since $\operatorname{im}(\varphi)$ is a subspace of $\mathbb{C}^k$). So

$$\dim V \le \dim W + \dim W^\perp.$$

On the other hand, since $W \cap W^\perp = \{0\}$, then $W \oplus W^\perp$ must be a *subspace* of V — more explicitly, we have a map $W \oplus W^\perp \to V$ given by $(w, u) \mapsto w + u$, and this map must have kernel $\{0\}$ since W and U only have 0 in common, so it must be injective and therefore it identifies $W \oplus W^\perp$ with a subspace of V.

So then we must have $\dim V \ge \dim W + \dim W^\perp$ as well. This implies equality holds in both statements, and therefore $W \oplus W^\perp = V$. □

> **Remark 5.32.** We used nondegeneracy only in the beginning, to show $W \cap W^\perp = \{0\}$. In particular, we don't need the form to be nondegenerate on V for this argument to work.

§5.3.2 Orthogonal Bases

Theorem 5.31 is quite powerful; one implication it has is the following.

> **Theorem 5.33**
>
> Given *any* symmetric bilinear form (for $\mathbb{R}$) or Hermitian form (for $\mathbb{C}$) on V, we can find an orthogonal basis for V — a basis $\{v_1, \dots, v_n\}$ such that $\langle v_i, v_j\rangle = 0$ for all $i \ne j$.

Concretely, an orthogonal basis is one where the matrix for $\langle -, -\rangle$ is diagonal (since all the entries not on the diagonal must be 0).

Proof. We use induction on $\dim V = n$.

To motivate the proof, we can use our geometric intuition — if our form makes sense geometrically, then we can take any vector u and the space $u^\perp$. Then by induction we can find an orthogonal basis for $u^\perp$, and combine this basis with u to get an orthogonal basis for V. This idea *almost* works in general, but we have to be careful, since being able to split V into a line and its orthogonal complement depends on nondegeneracy.

Case 1 (There exists some $u \in V$ such that $\langle u, u\rangle \ne 0$). Then let $W = \operatorname{Span}(u)$, which is a one-dimensional vector space. Since $\langle u, \rangle \ne 0$, the restriction of $\langle -, -\rangle$ to W is nondegenerate (since its corresponding matrix is the 1×1 matrix consisting of $\langle u, u\rangle$). So then by Theorem 5.31, we can write

$$V = W \oplus W^\perp,$$

where W is one-dimensional and $W^\perp$ is $(n-1)$-dimensional. By the inductive hypothesis $W^\perp$ has an orthogonal basis $\{v_2, \dots, v_n\}$, and adding in u gives an orthogonal basis for V.

Case 2 ($\langle v, v\rangle = 0$ for all $v \in V$). This is a very strong constraint, and we'll show it implies $\langle v, w\rangle = 0$ for any two vectors v and w. Then we're done, since *any* basis is an orthogonal basis.

Consider the equation

$$0 = \langle v + w, v + w\rangle = \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle = \langle v, w\rangle + \langle w, v\rangle.$$

In the case of $\mathbb{R}$, since $\langle -, -\rangle$ is symmetric, this immediately implies $\langle v, w\rangle = 0$. Meanwhile in the case of $\mathbb{C}$, this implies $\langle v, w\rangle$ has a real part of 0, and we can perform the same argument with $v + iw$ instead to get that $\langle v, w\rangle$ has imaginary part 0 as well. □

In fact, we can slightly strenghten this statement.

Corollary 5.34

We can find an orthogonal basis $\{v_1, \dots, v_n\}$ for V such that $\langle v_i, v_i\rangle \in \{1, -1, 0\}$ for each i.

Proof. We can start with any orthogonal basis $\{x_1, \dots, x_n\}$ and simply scale it — scaling any vector x_i preserves orthogonality. For each i, if $\langle x_i, x_i\rangle = 0$ then we can take $v_i = x_i$. Otherwise we can take $v_i = cx_i$ for some real c which makes $\langle v_i, v_i\rangle = \pm 1$ — more explicitly, we take

$$v_i = \frac{1}{\sqrt{|\langle x_i, x_i\rangle|}} x_i.$$ □

This tells us that up to a choice of basis, there aren't actually that many possibilities for a symmetric bilinear or Hermitian form — we can always find a basis in which the form is a diagonal matrix with diagonal entries all ±1 and 0. In particular, $\langle -, -\rangle$ is nondegenerate if and only if the diagonal only consists of ±1 — if we had a 0 on the diagonal, then the matrix would not be invertible.

Using these matrices, we can also describe another useful property:

Definition 5.35. The form $\langle -, -\rangle$ is *positive definite* if $\langle v, v\rangle > 0$ for all nonzero vectors v.

In the orthogonal basis as described in Corollary 5.34, we can see that $\langle -, -\rangle$ is positive definite if and only if the diagonal only consists of $+1$ (since a -1 or 0 would correspond to a basis vector with $\langle v_i, v_i\rangle \le 0$). Conversely, if we have a basis for which the matrix only consists of $+1$'s, then in that basis the form is just the dot product or standard Hermitian form, and is therefore positive definite.

There are many different orthogonal bases which we could use to write $\langle -, -\rangle$ as a matrix with the described property. But interestingly, the way this matrix looks doesn't depend on the choice of basis!

Fact 5.36 (Sylvester's Law) — Given $\langle -, -\rangle$ on V, for *any* choice of orthogonal basis as described in Corollary 5.34, the number of 1's, -1's, and 0's on the diagonal are fixed.

These numbers are called the *signature* of $\langle -, -\rangle$. For example, the dot product written in any orthogonal basis will consist of only $+1$'s on the diagonal, and the pairing in Example 5.24 will always consist of one -1 and three $+1$'s.

Finally, we can also translate these results into ones about matrices, in the same way as we did for linear maps and operators, by starting off with a matrix A representing a form on $\mathbb{R}^n$ and changing basis.

Theorem 5.37

If A is a symmetric $n \times n$ matrix over $\mathbb{R}$, then there exists a matrix $P \in \mathrm{GL}_n(\mathbb{R})$ such that $P^\intercal AP$ is diagonal, and all its diagonal entries are 1, -1, or 0. Furthermore, A is positive definite if and only if $A = Q^\intercal Q$ for an invertible matrix Q.

The first statement follows directly from our observation earlier that changing the basis corresponds to replacing A with $P^\intercal AP$. To see the second, we've seen that the form corresponding to A is positive definite if and only if the rewritten matrix $P^\intercal AP$ is actually I. We can let $Q = P^{-1}$ and multiply by $Q^\intercal$ and Q on the two sides, to get that this occurs if and only if $A = Q^\intercal Q$.

An analogous statement is true for matrices over $\mathbb{C}$ instead (using Hermitian matrices and taking the adjoints instead of transposes).

§5.3.3 Orthogonal Projection

Suppose we have a vector space V over $\mathbb{R}$ or $\mathbb{C}$ with a symmetric bilinear form or Hermitian form $\langle -, - \rangle$, and a subspace $W \subset V$ for which $\langle -, - \rangle$ restricted to W is nondegenerate. Then Theorem 5.31 tells us that $V = W \oplus W^\perp$, so every vector $v \in V$ can be written uniquely as a sum $w + u$ with $w \in W$ and $u \in W^\perp$.

Question 5.38. How can we compute w and u?

In a geometric setting, computing w and u corresponds to splitting v into a piece which lies in W, and another piece perpendicular to W.

Definition 5.39. The *orthogonal projection* $\pi\colon V \to W$ is the linear map sending $v \mapsto w$.

Note that by definition, $v - \pi(v) \perp W$.

Orthogonal projection is really useful. In geometric situations, w is the vector in W closest to v. So calculating w given v comes up a lot in data analysis, especially in using *least squares approximation*.

It turns out that there's a nice way of finding the map π assuming that we have an orthogonal basis for W. (If we don't already have an orthogonal basis for W, then we'd start by finding one — next class we'll discuss how to do this in the special case where the form is positive definite.) Let this orthogonal basis be $\{w_1, \dots, w_i\}$; then since the form is nondegenerate on W, we have $\langle w_i, w_i \rangle \neq 0$ for all i.

Now take our vector $v \in V$, so we want to find the coefficients c_i for which

$$v = c_1 w_1 + \cdots + c_k w_k + u$$

with $u \perp W$. To find these coefficients, we can simply pair v with the basis vectors — if v is in this form, then we have

$$\langle w_1, v \rangle = c_1 \langle w_1, w_1 \rangle + c_2 \langle w_1, w_2 \rangle + \cdots = c_1 \langle w_1, w_1 \rangle,$$

since w_1 is orthogonal to all the other w_i (since they form an orthogonal basis) and to u (since $u \in W^\perp$). The same occurs for all other indices, so we get

$$c_i = \frac{\langle w_i, v \rangle}{\langle w_i, w_i \rangle}$$

for each i, and then $\pi(v) = c_1 w_1 + \cdots + c_k w_k$ for these values of c_i.

Example 5.40

Consider the vector space $\mathbb{R}^3$ with the dot product, and let W be the span of $w_1 = (1, 1, 1)^\intercal$ and $w_2 = (1, 1, -2)^\intercal$ (which form an orthogonal basis of W). Find the projection of $v = (1, 2, 3)^\intercal$ onto W.

Solution. We can compute $\langle w_1, w_1 \rangle = 3$, $\langle w_1, w_2 \rangle = 6$, $\langle w_1, v \rangle = 6$, and $\langle w_2, v \rangle = -3$. So then we have

$$\pi(v) = \frac{6}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 3/2 \\ 3 \end{bmatrix}. \qquad \square$$

One interesting application of this formula is that if we take W to be the entire vector space V, then this tells us how to find the coordinates of any vector v with respect to a given orthonormal basis of V.

§5.4 Euclidean and Hermitian Spaces

Definition 5.41. A *Euclidean space* is a vector space over $\mathbb{R}$ equipped with a *positive definite* symmetric bilinear form. Similarly, a *Hermitian space* is a vector space over $\mathbb{C}$ equipped with a positive definite Hermitian form.

We've seen already (in Corollary 5.34) that if V is Euclidean or Hermitian, then it has an *orthonormal basis*:

Definition 5.42. An *orthonormal basis* is a basis $\{v_1, \ldots, v_n\}$ such that

$$\langle v_i, v_j \rangle = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

This is because Corollary 5.34 implies we can find an orthogonal basis in which each $\langle v_i, v_i \rangle$ is ± 1 or 0, and then positive definiteness guarantees that they are all 1.

Many of the results we proved earlier rely on nondegeneracy. It turns out that nondegeneracy always holds in this situation:

Proposition 5.43

If V with $\langle -, - \rangle$ is Euclidean or Hermitian, then $\langle -, - \rangle$ restricted to *any* subspace W is nondegenerate.

Proof. Recall that by definition, $\langle -, - \rangle$ restricted to W is degenerate if and only if there are nonzero vectors $w \in W$ which are orthogonal to all of W. But each vector w is not orthogonal to *itself*, since $\langle w, w \rangle > 0$. □

So then everything we've proved regarding splitting $V = W \oplus W^\perp$ and calculating orthogonal projections does hold in any Euclidean or Hermitian space.

§5.4.1 The Gram–Schmidt Algorithm

Question 5.44. Given a Euclidean or Hermitian space, how can we find an orthonormal basis?

Suppose we have a Euclidean or Hermitian space V, and we start off with *some* basis $\{v_1, \ldots, v_n\}$. We'd like to turn this basis into a basis $\{u_1, \ldots, u_n\}$ which is orthonormal.

We'll inductively build this basis, by going through our original basis and correcting it one vector at a time. More precisely, let $V_k = \text{Span}(v_1, \ldots, v_k)$ for each k. Then we'll construct $u_1, u_2, \ldots$ so that $\text{Span}(u_1, \ldots, u_k)$ is *also* V_k for all k, and the u_i form an orthonormal basis for V_k.

Step 0. First we'll find u_1. We can "fix" v_1 by simply scaling it — we take

$$u_1 = \frac{1}{\sqrt{\langle v_1, v_1 \rangle}} v_1,$$

which produces an orthonormal basis for V_1.

Step 1. We now want to find an orthonormal basis for V_2, building off the one we have for V_1. The problem with our original basis is that v_2 may not be orthogonal to v_1. So we first set $x_2 = \text{proj}_{V_1} v_2$ (this denotes the projection of v_2 onto V_1), and $y_2 = v_2 - x_2$. Then y_2 is orthogonal to all of V_1 (and therefore to u_1), and $\text{Span}(u_1, y_2) = V_2$. So finally, we can scale y_2 to get our basis vector

$$u_2 = \frac{1}{\sqrt{\langle y_2, y_2\rangle}} y_2.$$

We can essentially perform the same construction for all the following steps:

Step *k*. Suppose we've constructed an orthonormal basis $\{u_1, \dots, u_k\}$ for V_k, and we now want to find u_{k+1}. First set $x_{k+1} = \text{proj}_{V_k} v_{k+1}$ and $y_{k+1} = v_{k+1} - x_{k+1}$ — so y_{k+1} is essentially the part of v_{k+1} which is orthogonal to V_k. Note that replacing v_{k+1} with y_{k+1} doesn't change the span of our first $k+1$ vectors, since x_{k+1} is in V_k; also note that y_{k+1} is nonzero, since v_{k+1} cannot be in V_k. Then we can scale again and take our new basis vector to be

$$u_{k+1} = \frac{1}{\sqrt{\langle y_{k+1}, y_{k+1}\rangle}} y_{k+1}.$$

Repeating this process eventually produces an orthonormal basis for the entire space V.

What's nice about this algorithm is that when we're attempting to project v_{k+1} onto V_k, we *already* have an orthonormal basis for V_k, which means we know how to do the projection — we have

$$\text{proj}_{V_k} v_{k+1} = \sum \frac{\langle u_i, v_{k+1}\rangle}{\langle u_i, u_i\rangle} u_i = \sum \langle u_i, v_{k+1}\rangle u_i.$$

So this algorithm is quite computationally feasible.

We can also rewrite the result of the Gram–Schmidt algorithm in terms of matrices. Suppose we start off with a matrix $M \in \text{GL}_n(\mathbb{R})$, so we can think of the columns of M as a basis $\{v_1, \dots, v_n\}$ for $\mathbb{R}^n$. The algorithm tells us that we can correct this basis to turn it into an orthonormal basis $\{u_1, \dots, u_n\}$ with respect to the dot product. More specifically, it tells us that we can correct the basis in a way such that $\text{Span}(u_1, \dots, u_k) = \text{Span}(v_1, \dots, v_k)$ for all k — so then we have $u_1 = a_{11}v_1$, $u_2 = a_{12}v_1 + a_{22}v_2$, $u_3 = a_{13}v_1 + a_{23}v_2 + a_{33}v_3$, and so on, for some coefficients a_{ij}. We can rewrite this as

$$\begin{bmatrix} | & | & \cdots & | \\ u_1 & u_2 & \cdots & u_k \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{12} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} | & | & \cdots & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & \cdots & | \end{bmatrix}.$$

So this gives the following statement in terms of matrices:

Proposition 5.45

Given any $M \in \text{GL}_n(\mathbb{R})$, there exists a lower triangular matrix Q and an orthogonal matrix R such that $M = QR$.

Here R is the matrix of the u_i (corresponding to the new basis), and Q is the inverse of the matrix (a_{ij}). This result can be computationally useful.

§5.5 The Spectral Theorem

We'll now work specifically over $\mathbb{C}$ — let V be a Hermitian space with pairing $\langle -, - \rangle$.

§5.5.1 Some Definitions

Definition 5.46. Given a linear operator $T\colon V \to V$, its *adjoint* is the linear operator $T^*\colon V \to V$ defined as follows: Take an orthonormal basis $\{u_1, \dots, u_n\}$ of V. Then if the operator T corresponds to the matrix M, its adjoint T^* is the linear operator corresponding to the matrix M^*.

So we're essentially using our definition of the adjoint for *matrices*, working over $\mathbb{C}^n$ with the standard Hermitian form, to define the adjoint for *linear operators* in abstract Hermitian spaces (since choosing an orthonormal basis maps $V \to \mathbb{C}^n$ and $\langle -, - \rangle$ to the standard Hermitian form).

This definition is quite weird, because it asks us to pick an orthonormal basis — if we picked a different basis, would we still get the same operator T^*? Fortunately, the answer is yes — it's possible to describe T^* without referencing a basis at all.

Proposition 5.47

The operator T^* has the property that $\langle Tv, w\rangle = \langle v, T^*w\rangle$ for all v and w.

Proof. We can translate this statement to one about $\mathbb{C}^n$ by picking an orthonormal basis $\{u_1, \dots, u_n\}$. Suppose that in this basis, v corresponds to a column vector $\vec{x}$ and w to $\vec{y}$. Then Tv corresponds to $M\vec{x}$, so using properties of the adjoint, we have

$$\langle Tv, w\rangle = (M\vec{x})^*\vec{y} = \vec{x}^*M^*\vec{y} = \vec{x}^*(M^*\vec{y}) = \langle v, T^*w\rangle,$$

since we defined T^* to correspond to M^*. □

This *uniquely* determines the linear operator — if we take an orthonormal basis $\{u_1, \dots, u_n\}$ and set $v = u_i$ and $w = u_j$, then $\langle u_i, T^*u_j\rangle$ gives the ith coordinate of T^*u_j, so over all i and j this determines the value of T^* on each basis vector, and therefore on all of V. So this means the definition of T^* is independent of the choice of basis.

Now that we have the concept of an adjoint in an abstract Hermitian space, we can take some of the definitions we had in $\mathbb{C}^n$ that referenced adjoints, and move them to our Hermitian space as well.

Definition 5.48. A linear operator $T\colon V \to V$ is a *Hermitian operator* if $T^* = T$.

If we fix an orthonormal basis, then a Hermitian operator corresponds to a Hermitian matrix (one with the property that $A^* = A$). Note that this condition is equivalent to stating that for all v and w we have

$$\langle Tv, w\rangle = \langle v, Tw\rangle.$$

Definition 5.49. A linear operator $T\colon V \to V$ is a *unitary operator* if T^*T is the identity operator.

Equivalently, if we fix an orthonormal basis, then a unitary operator corresponds to a unitary matrix (a matrix such that $U^*U = I$). We've seen that unitary *matrices* are the matrices which preserve the standard Hermitian product, and the same is true here — T is Hermitian if and only if for all v and w we have

$$\langle Tv, Tw\rangle = \langle v, w\rangle.$$

There is a more general property that encapsulates both of these:

Definition 5.50. A linear operator $T\colon V \to V$ is *normal* if $TT^* = T^*T$.

Hermitian operators are normal because $T^* = T$, and unitary matrices are normal because $T^* = T^{-1}$ — any matrix commutes with itself and with its inverse. But this property is more general than being either Hermitian or unitary — there exist operators which are normal but neither Hermitian nor unitary.

Example 5.51

Consider the matrix

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

This is normal, but not Hermitian or unitary — A^* isn't A or A^{-1}, but it does commute with A.

We can also rewrite this condition in terms of the pairing — if T is normal, then for all v and w we have

$$\langle Tv, Tw\rangle = \langle v, T^*Tw\rangle = \langle v, TT^*w\rangle = \langle T^*v, T^*w\rangle.$$

The converse is true as well (since if $\langle v, T^*Tw\rangle = \langle v, TT^*w\rangle$ for all v and w, we must have $T^*T = TT^*$).

§5.5.2 The Spectral Theorem

Theorem 5.52 (Spectral Theorem)

Let V be a Hermitian space, and let $T\colon V \to V$ be a *normal* linear operator. Then V has an orthonormal basis $\{u_1, \dots, u_n\}$ where each u_i is an eigenvector of T.

We've previously discussed how to diagonalize a linear operator T, and how to find an orthonormal basis for a Hermitian space V. But the Spectral Theorem tells us that we can do *both* at once — we can find a basis that answers both questions simultaneously. In particular, it means a normal linear operator is *always* diagonalizable — we don't need to use the more complicated Jordan normal form.

We can rewrite the Spectral Theorem in terms of matrices as well — assume that V is $\mathbb{C}^n$ under the standard Hermitian product. Then the theorem tells us that given a matrix M such that $M^*M = MM^*$, we can find an orthonormal eigenbasis — this means we can find a *unitary* matrix P such that $P^{-1}MP$ is diagonal (here the columns of P are the eigenbasis, and P is unitary since its columns are orthonormal). Note that since P is unitary, we also have $P^{-1}MP = P^*MP$.

In this section we've been working over $\mathbb{C}$ because the theorem at this level of generality is false over $\mathbb{R}$. There *is* a version of the Spectral Theorem over $\mathbb{R}$ — given a Euclidean space V and a *symmetric* linear operator $T\colon V \to V$, we can find an orthonormal eigenbasis.

But this isn't true in as much generality as it is for $\mathbb{C}$ (we can think of symmetric as the analog of Hermitian). For example, we can't even necessarily diagonalize an orthogonal matrix in the first place (we've seen that rotations in $\mathbb{R}^2$ are orthogonal, and they have no eigenvectors).

Example 5.53

Consider the symmetric matrix

$$M = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}.$$

Then the vectors $\frac{1}{\sqrt{2}}(1,1)^\intercal$ and $\frac{1}{\sqrt{2}}(1,-1)^\intercal$ form an orthonormal eigenbasis, with eigenvalues 2 and 4.

Now we'll prove the Spectral Theorem. First we'll prove some useful properties of the adjoint.

Lemma 5.54

Let $T: V \to V$ be a linear operator, and let W be a subspace of V such that T preserves W. Then T^* preserves $W^\perp$.

Recall that T preserving W means that $T(W) \subset W$ — or in other words, for all $w \in W$, we have $Tw \in W$ as well.

Proof. It suffices to show that for any $u \in W^\perp$, we have $T^*u \in W^\perp$ as well, meaning that $\langle w, T^*u\rangle = 0$ for all $w \in W$. But by Proposition 5.47, we have

$$\langle w, T^*u\rangle = \langle Tw, u\rangle.$$

Since T preserves W, we have $Tw \in W$ and $u \in W^\perp$, which means this pairing must be 0. So $\langle w, T^*u\rangle = 0$ for all $w \in W$, and therefore T^*u is in $W^\perp$. □

Lemma 5.55

Let T be a *normal* linear operator. If $Tv = \lambda v$, then $T^*v = \overline{\lambda}v$.

In other words, this states that T and T^* have the same set of eigenvectors, and their eigenvalues are complex conjugates.

Proof. First we'll solve the specific case where $\lambda = 0$. Then we know v is in $\ker(T)$, and we want to show v is in $\ker(T^*)$ as well. Since $Tv = 0$ and T is normal, we have

$$\langle T^*v, T^*v\rangle = \langle Tv, Tv\rangle = 0.$$

But since V is Hermitian, $\langle -, -\rangle$ is positive definite, so we must have $T^*v = 0$.

Now for the general case, let $S = T - \lambda I$. Then v is in $\ker(S)$, and meanwhile we have $S^* = T^* - \overline{\lambda}I$. We can check that S is still normal — S and S^* commute because T and T^* do. So then by the special case shown earlier, we have that v is in $\ker(S^*)$ as well, which means $T^*v = \overline{\lambda}v$. □

Remark 5.56. The main idea here, of studying a general eigenvector by shifting its eigenvalue to 0, is one that came up quite frequently when we studied eigenvectors previously.

Now with this, we can prove the Spectral Theorem.

Proof of Theorem. We'll use induction on the dimension of V — we'll break V as a direct sum of two smaller pieces, and inductively find an orthonormal eigenbasis of each piece.

Since we're working over $\mathbb{C}$, we know we can find at least one eigenvector $w \in V$ and a corresponding eigenvalue λ, so that $Tw = \lambda w$. Since $\langle w, w\rangle > 0$, we can scale w such that $\langle w, w\rangle = 1$.

Now let $W = \text{Span}(w)$, and split $V = W \oplus W^\perp$ (we can do this by Theorem 5.31 since $\langle -, -\rangle$ is positive definite and therefore nondegenerate). We know T preserves W (since T scales w), so in order to be able to induct, we want to show that T preserves $W^\perp$ as well.

We know that $Tw = \lambda w$, so by Lemma 5.55 (since T is normal), w is an eigenvector for T^* as well. So then T^* *also* preserves W, and by Lemma 5.54 this means $(T^*)^* = T$ preserves $W^\perp$.

So now we can split $V = W \oplus W^\perp$. Since T preserves both W and $W^\perp$, it acts separately on the two pieces. So by the inductive hypothesis, we can find an orthonormal eigenbasis for T acting on $W^\perp$; then adding w to this list gives an orthonormal eigenbasis for V. □

In the case of Euclidean spaces over $\mathbb{R}$, most of the argument still works as written. The one part which *doesn't* work is the beginning, where we find an eigenvector and an eigenvalue — the argument breaks if we can't find *any* real eigenvectors. This is why the Spectral Theorem *does* hold for symmetric matrices — we showed in Proposition 5.21 that for any symmetric matrix (or more generally any *Hermitian* matrix), all its eigenvalues are real. So we can always find one eigenvector w and eigenvalue λ to get started, and the rest of the argument works in the exact same way.

Remark 5.57. In fact, in the case of symmetric matrices over $\mathbb{R}$, there's another proof of the Spectral Theorem via Lagramge multipliers to find w and λ — the value of λ which comes from using Lagrange multipliers is actually an eigenvalue.

§5.5.3 Application to Quadratic Forms

Suppose we have a quadratic form $f(x, y) = ax^2 + bxy + cy^2$ (a function in x and y which only contains terms of degree 2). Then we can rewrite f in terms of a symmetric matrix — if we take

$$M = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix},$$

then we have

$$f(x, y) = (x, y)M(x, y)^\intercal.$$

Example 5.58

The quadratic form $f(x, y) = 3x^2 - 2xy + 3y^2$ corresponds to the matrix

$$M = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}.$$

By the Spectral Theorem, there is an orthogonal change of coordinates — meaning we write $(x, y)^\intercal = P(x', y')^\intercal$ for an orthogonal matrix P — such that we have

$$P^\intercal MP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Then using this change of coordinates, we can rewrite

$$f(x, y) = \lambda_1(x')^2 + \lambda_2(y')^2,$$

which makes it much easier to understand what the original quadratic form does.

Example 5.59

For the form given in Example 5.58, we can set $x = \frac{1}{\sqrt{2}}(x' + y')$ and $y = \frac{1}{\sqrt{2}}(x' - y')$, and in the new coordinates we have

$$f(x, y) = 2(x')^2 + 4(y')^2.$$

We can do this with more than two variables as well — if we have a quadratic form $f(x_1, \ldots, x_n) = a_{11}x_{11}^2 + \cdots + a_{nn}x_{nn} + 2\sum_{i<j} a_{ij}x_ix_j$, then we can again choose an orthogonal matrix P such that $(x_1, \ldots, x_n)^\intercal = P(x_1', \ldots, x_n')^\intercal$, and our original quadratic form becomes $\lambda_1(x_1')^2 + \cdots + \lambda_n(x_n')^2$ — so we can eliminate all the cross-terms x_ix_j. More explicitly, we can obtain P by taking the symmetric matrix M of the a_{ij}'s and using the Spectral Theorem to find a new coordinate system (which is still orthonormal, and in which M becomes diagonal).

In two and three dimensions, this has a nice geometric interpretation as well:

Question 5.60. Consider a curve $ax^2 + bxy + cy^2 + dx + ey + f = 0$. What can this curve look like?

There's a few familar possibilities:

- An ellipse, for example given by $ax^2 + by^2 = 1$;
- A hyperbola, for example given by $ax^2 - by^2 = 1$;
- A parabola, for example given by $ax^2 - y = 0$;
- Two intersecting lines, given by $(a_1x + b_1y)(a_2x + b_2y) = 0$;
- Two parallel lines, for example given by $x^2 = a$;
- A single line, for example given by $x^2 = 0$;
- A single point, for example given by $x^2 + y^2 = 0$;
- The empty set, for example given by $x^2 + y^2 = -1$.

The first three cases — ellipses, hyperbolas, and parabolas — are called *conics*; the remaining cases are all degenerate.

Theorem 5.61

After an isometry, all curves $ax^2 + bxy + cy^2 + dx + ey + f = 0$ look like one of the curves on this list.

Proof. Let $\vec{v} = (x, y)^\intercal$, so then we can rewrite the equation as

$$\vec{v}^\intercal A\vec{v} + B\vec{v} + f = 0$$

for matrices A and B. First we'll deal with the quadratic part using the Spectral Theorem — we can find an orthogonal change of basis in which A becomes diagonal, so then our equation becomes

$$\lambda_1x^2 + \lambda_2y^2 + b_1x + b_2y + f = 0$$

(note that we're now using x and y to refer to the new variables). First, if λ_1 and λ_2 are nonzero, then we can complete the square to get rid of the linear terms — if we send $x \mapsto x + b_1/2\lambda_1$ and $y \mapsto y + b_2/\lambda_2$, then we get an equation of the form

$$\lambda_1x^2 + \lambda_2y^2 = c.$$

If c is nonzero then we either get an ellipse, hyperbola, or the empty set; if $c = 0$ then we either get one point or two intersecting lines (depending on whether λ_1 and λ_2 have the same or opposite sign).

We can perform a similar analysis when one of λ_1 and λ_2 is 0, and this will give the remaining cases — for instance, if $\lambda_1 = 0$ and $\lambda_2 \neq 0$, then we get a parabola. □

The importance of the fact that our new basis is *orthogonal* is that the transformations we perform on x and y are *isometries* — we've essentially just rotated (or reflected) the coordinate axes. So our original curve has the same shape as the new one, and if we want to return to the original coordinate system, we can simply rotate back.

§6 Linear Groups

§6.1 Introduction

So far, we've been studying group theory and linear algebra. We'll now study a topic related to both — we'll look at groups of *matrices* with certain properties. The group of *all* invertible matrices, $\mathrm{GL}_n(\mathbb{R})$, has several interesting subgroups: the special linear group $\mathrm{SL}_n(\mathbb{R})$, the matrices with determinant 1; the orthogonal group O_n, consisting of matrices such that $A^\intercal = A^{-1}$; and the special orthogonal group SO_n, which is their intersection. All of these are groups of matrices which preserve some linear algebraic property — $\mathrm{SL}_n(\mathbb{R})$ preserves volume, and O_n preserves the dot product.

We can also work over $\mathbb{C}$ instead of $\mathbb{R}$, and consider subgroups of $\mathrm{GL}_n(\mathbb{C})$. One interesting subgroup is still $\mathrm{SL}_n(\mathbb{C})$. Another is the *unitary* group U_n, consisting of matrices such that $A^* = A^{-1}$ — similarly to how O_n preserves the dot product, U_n preserves the standard Hermitian product. We also have their intersection, the *special unitary group* SU_2, which consists of matrices which both have determinant 1 and are unitary.

These are all examples we've seen before, but there are many others. We could also look at matrices preserving *other* bilinear forms — for example, the form defined by

$$I_{p,q} = \left[\begin{array}{c|c} I_p & 0 \\ \hline 0 & -I_q \end{array}\right] = \left[\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & & & \\ 0 & 1 & \cdots & 0 & & 0 & \\ \vdots & \vdots & \ddots & \vdots & & & \\ 0 & 0 & \cdots & 1 & & & \\ \hline & & & & -1 & \cdots & 0 \\ & 0 & & & \vdots & \ddots & \vdots \\ & & & & 0 & \cdots & -1 \end{array}\right].$$

These matrices would satisfy the equation $A^\intercal I_{p,q} A = I_{p,q}$; they form another interesting subgroup of $\mathrm{GL}_n(\mathbb{R})$.

What's special about working with $\mathbb{R}$ or $\mathbb{C}$ (as opposed to a finite field) is that they have a notion of distance. We can think of matrices in $\mathrm{GL}_n(\mathbb{R})$ as a subset of $\mathbb{R}^{n^2}$, so then we have a way of measuring the distance between two matrices — this means subgroups $G \leq \mathrm{GL}_n(\mathbb{R})$ inherit a *metric.* The same is true for $G \leq \mathrm{GL}_n(\mathbb{C})$, since we also have a definition of distance between complex numbers. This means that the group structure and topology can interact.

In particular, note that in all these cases, the group operation of multiplication of matrix multiplication is *continuous*, and so is its inverse. Previously when studying groups, we looked at homomorphisms between them, which preserved the group structure. Now we have both a group structure *and* a topological structure, so we'll look at *continuous* homomorphisms.

Example 6.1

There is a continuous homomorphism $(\mathbb{R}, +) \to \mathrm{SO}_2$, given by sending $\theta \mapsto \rho_\theta$. We can think of this homomorphism geometrically — since SO_2 consists exactly of rotation matrices, we can think of it as $\mathbb{R}/2\pi\mathbb{Z}$, or as a circle. Then this homomorphism maps the line $\mathbb{R}$ to the circle by wrapping it around the circle infinitely many times.

Example 6.2

We can think of O_2 geometrically as *two* circles — one circle represents SO_2, and the other represents the set of reflections.

Both these examples are one-dimensional, but we'll soon see an example of a group which geometrically is a higher-dimensional figure, and where this geometric intuition is really useful in understanding the group.

§6.2 The Geometry of SU_2

Recall that SU_2 is the set of 2×2 matrices over $\mathbb{C}$ which have determinant 1 and are unitary (meaning $A^* = A^{-1}$). We'd like to figure out what SU_2 "looks like."

§6.2.1 An Explicit Description

To start with, we'd like to get a more explicit description for SU_2. Suppose we have a matrix

$$A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \in \mathrm{SU}_2.$$

Then we can calculate

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix} = \begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix}$$

(since A is supposed to have determinant 1), while

$$A^* = \begin{bmatrix} \overline{\alpha} & \overline{\gamma} \\ \overline{\beta} & \overline{\delta} \end{bmatrix}.$$

Setting these equal gives that we must have

$$A = \begin{bmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{bmatrix},$$

and then the condition $\det(A) = 1$ becomes $|\alpha|^2 + |\beta|^2 = 1$.

We can now write $\alpha = x_0 + ix_1$ and $\beta = x_2 + ix_3$, so then we have

$$A = x_0 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + x_1 \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} + x_2 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

with $x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1$. The first matrix is I, and we can name the others $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$. These satisfy the properties that $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -I$, $\mathbf{ij} = \mathbf{k}$ and its cyclic variants, and $\mathbf{ji} = -\mathbf{k}$ and its cyclic variants. This actually gives rise to an interesting structure:

> **Definition 6.3.** The *quaternions*, denoted by $\mathbb{H}$, are the group of elements $x_0 I + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ for $x_0, x_1, x_2, x_3 \in \mathbb{R}$ under multiplication.

Since we know how to multiply any two of these matrices, $\mathbb{H}$ is closed under multiplication, so it is a valid group — we can think of it as a four-dimensional version of $\mathbb{C}$, except that multiplication isn't commutative. But what's important for our purposes is that $\mathbb{H}$ is a four-dimensional vector space over $\mathbb{R}$, and SU_2 is a subset of this four-dimensional vector space. More specifically, it's the subset satisfying the equation

$$x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1,$$

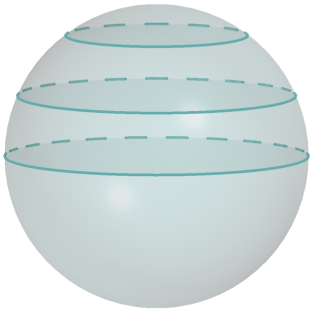
which defines a three-dimensional sphere (in a four-dimensional space). So geometrically, SU_2 is the 3-sphere.

§6.2.2 Geometry of a Sphere

The 3-sphere is hard to picture, so we'll start by thinking about the 2-sphere S^2 — the set of solutions in $\mathbb{R}^3$ to $x_0^2 + x_1^2 + x_2^2 = 1$. One natural way to think of points on S^2 is in terms of latitude and longitude.

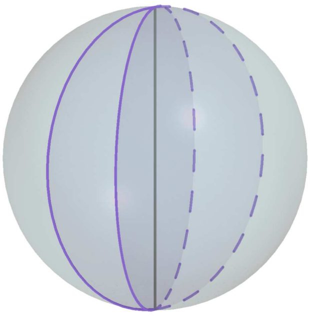
Definition 6.4. The *latitudes* of a sphere are the sets of points on the sphere with fixed x_0; we use Lat_c to denote the set of points with $x_0 = c$.

Geometrically, the latitudes of the 2-sphere come from taking horizontal slices of the sphere.



So the latitudes of the 2-sphere are all circles — if we start at the top of the sphere and move down, then the latitudes start at a point, grow larger until the equator of the sphere, and then shrink back to a point.

Definition 6.5. The *longitudes* of a sphere are the circles of radius 1 which pass through the north and south poles.



So longitudes correspond to slicing the sphere at different angles.

Now we can do the same for the 3-sphere. The latitudes are still defined as

$$\mathrm{Lat}_c = \{(x_0, x_1, x_2, x_3) \mid x_0 = c\} \cap S^3,$$

the set of points with fixed first coordinate. Note that now the latitudes are 2-spheres. The latitudes at $c = \pm 1$ are single points, and the latitude at $c = 0$ is the largest sphere; this is called the *equator*, and denoted $\mathbb{E}$.

We can also consider the longitudes of the 3-sphere. We'll define these more precisely later, but they will still be circles of radius 1 passing through the north and south poles $(\pm 1, 0, 0, 0)$.

We can use latitudes and longitudes to describe our 3-sphere — every point lies on a unique latitude, and every point except the north and south poles lies on a unique longitude. Meanwhile, each latitude and longitude intersect at two points. We'll now see how this description of S^3 in terms of latitudes and longitudes can be used to understand the group structure of SU_2.

§6.2.3 Latitudes of SU_2

Theorem 6.6

The conjugacy classes of SU_2 are precisely the latitudes.

This has a useful corollary — recall that an element is in the *center* of a group if and only if its conjugacy class has size 1. But the only latitudes which consist of just one element are the latitudes at ±1, so this implies that the center $Z(\mathrm{SU}_2)$ is precisely $\pm I$.

Proof. The key observation is that I has trace 2, while $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ all have trace 1. So then if a matrix $A \in \mathrm{SU}_2$ has coordinates (x_0, x_1, x_2, x_3) on the sphere, we have $\mathrm{tr}(A) = 2x_0$. So then latitudes correspond exactly to slices of SU_2 with fixed trace.

So it suffices to show that two matrices A and A' in SU_2 are conjugate if and only if they have the same trace. One direction is clear — trace is preserved by conjugation, so if A and A' are conjugate, then they *must* have the same trace.

For the other direction, we want to show that if $\mathrm{tr}(A) = \mathrm{tr}(A')$, then A and A' are conjugate to each other — meaning that $A' = P^{-1}AP$ for some $P \in \mathrm{SU}_2$. First note that A and A' both have characteristic polynomial

$$t^2 - t \cdot \mathrm{tr}(A) + 1,$$

since their traces are $\mathrm{tr}(A)$ and their determinants are 1. Let λ_1 and λ_2 be the roots of this polynomial. Then by the Spectral Theorem, we can diagonalize A using an orthonormal basis, which means we can find a *unitary* matrix Q such that

$$Q^{-1}AQ = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

But then we can do the same for A'. This immediately implies that A and A' are conjugate to the same matrix in U_2, and therefore to each other.

But we actually need to show that they're conjugate in SU_2. It turns out this isn't much harder — we can scale the matrix Q described above to have determinant 1. More explicitly, suppose Q is a unitary matrix, with $\det Q = \delta$. Then since Q is unitary, we have $Q^*Q = I$, so

$$1 = \det(Q^*)\det(Q) = \overline{\delta}\delta,$$

which means $|\delta| = 1$. Now scale Q to the matrix $\widetilde{Q} = \gamma Q$ for one of the two complex numbers γ such that $\gamma^2 = \delta^{-1}$. Then $\widetilde{Q}$ is still unitary, as $\widetilde{Q}^*\widetilde{Q} = \overline{\gamma}Q^* \cdot \gamma Q = Q^*Q = I$ (since δ has magnitude 1, so γ must have magnitude 1 as well). But now $\widetilde{Q}$ also has determinant 1.

So then A and A' are both conjugate to the described diagonal matrix in SU_2 as well, and therefore they are in the same conjugacy class. □

Remark 6.7. When studying conjugacy classes of *finite* groups, we had the class equation

$$|G| = \sum_i |C_i|,$$

where the C_i denote the conjugacy classes of G. Here SU_2 is infinite, so it doesn't make sense to discuss the size of sets. But it *does* make sense to discuss volume, so the analog of the class equation here is

$$\mathrm{Vol}(\mathrm{SU}_2) = \int_{-1}^{1} \mathrm{Vol}(\mathrm{Conj}_c)\, dc.$$

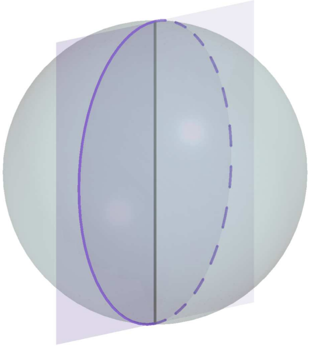
This idea is quite useful when studying SU_2 more deeply — if we want to integrate a geometric quantity over the entire group, we can first integrate over each *conjugacy class.*

§6.2.4 Longitudes of SU_2

When discussing the geometry of the 2-sphere, we also discussed *longitudes*. We'll now define the longitudes of S^3 more precisely:

> **Definition 6.8.** Given a point $x \in \mathbb{E}$, its corresponding longitude, denoted Long_x, is the circle passing through the north pole, the south pole, and x.

Alternatively, we have $\mathrm{Long}_x = \mathrm{Span}(I, x) \cap S^3$ — here $\mathrm{Span}(I, x)$ is the plane passing through the north and south pole and x, and intersecting it with the sphere gives the unit circle of this plane.



Theorem 6.9

For each $x \in \mathbb{E}$, the longitude Long_x is a subgroup of SU_2.

Proof. We'll first prove this for the specific case $x = \mathbf{i}$. Suppose we have two points $cI + s\mathbf{i}$ and $c'I + s'\mathbf{i}$ in $\mathrm{Long}_{\mathbf{i}}$. Then since $\mathbf{i}^2 = -1$, their product is still a linear combination of I and $\mathbf{i}$, and is therefore in $\mathrm{Span}(I, \mathbf{i})$ as well. But their product is also in SU_2 since SU_2 is a group, so it's in $\mathrm{Long}_{\mathbf{i}}$ as well. This means $\mathrm{Long}_{\mathbf{i}}$ is closed under multiplication, so it's a group (the identity is I, and the inverse of $c + s\mathbf{i}$ is $c - s\mathbf{i}$).

But now if we take *any* point $x \in \mathbb{E}$, we know x is conjugate to $\mathbf{i}$ (since the conjugacy classes are precisely the latitudes). Then Long_x is conjugate to $\mathrm{Long}_{\mathbf{i}}$ as well, so it's also a subgroup. □

This proof shows that not only are the longitudes all subgroups, but they're also conjugate to each other.

In fact, we have an isomorphism from the circle group $\mathbb{R}/2\pi\mathbb{Z}$ to Long_x given by $\theta \mapsto \cos\theta \cdot I + \sin\theta \cdot x$ (this can be shown in the same way — it's true for $x = \mathbf{i}$ by straightforward computation, and we can extend the result to all $x \in \mathbb{E}$ using the fact that $\mathbb{E}$ is a conjugacy class).

The longitudes have other applications as well:

> **Fact 6.10** — For any $x \in \mathbb{E}$, the centralizer $Z(x)$ is exactly Long_x.

It's clear that Long_x must be *contained* in $Z(x)$ — it's a subgroup of SU_2 containing x which is abelian (since the circle group $\mathbb{R}/2\pi\mathbb{Z}$ is abelian), so all its elements must commute with x. But it turns out that equality holds.

§6.2.5 Connection to SO_3

In Proposition 4.20, we saw that in any group action, given some $s \in S$ there is a bijection between left cosets of the stabilizer of s and elements of the orbit of s. In particular, taking the action to be conjugation, for each $g \in G$, there is a bijection between the left cosets of $Z(g)$ and elements of $C(g)$. Here, if we take g to be $\mathbf{i}$, then we get a bijection between its conjugacy class $\mathbb{E}$ and the cosets of $\text{Long}_{\mathbf{i}}$, which are all circles (not necessarily through the north and south poles). We can actually take this further using group theoretic terms.

We know SU_2 acts on $\mathbb{E}$ by conjugation, since $\mathbb{E}$ is a conjugacy class. In fact, conjugation by each element of SU_2 defines a *linear operator* on the subspace of $\mathbb{H}$ with $x_0 = 0$ (since conjugation preserves trace, and therefore x_0 — the linearity of this operator follows directly from the distributivity of matrix multiplication). So this defines a group homomorphism $\rho\colon \text{SU}_2 \to \text{GL}_3(\mathbb{R})$, where for each $g \in \text{SU}_2$, we define $\rho(g)$ as the linear operator on $\mathbb{H}$ given by $v \mapsto gvg^{-1}$.

But we know that this linear operator preserves $\mathbb{E}$, which is the unit sphere inside $\mathbb{H}$. So this means $\rho(g)$ preserves the length of unit vectors, and therefore the length of *all* vectors — so it's actually an isometry, and our homomorphism is actually a homomorphism $\rho\colon \text{SU}_2 \to \text{O}_3$.

But we can actually say even more. We know all orthogonal matrices have determinant 1 or -1. But SU_2 is connected (we can start at any point and reach any other point by taking a continuous path), and ρ is continuous, so $\det(\rho(g))$ cannot ever jump from 1 to -1. This means $\det(\rho(g))$ is constant over all g; and since the determinant of the identity is 1, this means $\det(\rho(g)) = 1$ for all g. So then ρ is actually a homomorphism $\text{SU}_2 \to \text{SO}_3$.

In fact, it's possible to show that as the homomorphism $\rho\colon \text{SU}_2 \to \text{SO}_3$ is surjective, and its kernel is $\{\pm I\}$. So this gives the following result:

Fact 6.11 — The quotient $\text{SU}_2/\{\pm I\}$ is isomorphic to SO_3.

§6.3 One-Parameter Groups

Definition 6.12. A *one-parameter group* in $\text{GL}_n(\mathbb{R})$ or $\text{GL}_n(\mathbb{C})$ is a differentiable homomorphism $\varphi\colon \mathbb{R} \to \text{GL}_n(\mathbb{R})$ or $\varphi\colon \mathbb{R} \to \text{GL}_n(\mathbb{C})$.

This means we have a function φ from $\mathbb{R}$ to $\text{GL}_n(\mathbb{R})$ or $\text{GL}_n(\mathbb{C})$ such that $\varphi(s+t) = \varphi(s)\varphi(t)$ for all real s and t, and if we consider our matrices as subsets of $\mathbb{R}^{n^2}$ or $\mathbb{R}^{2n^2}$ (by taking the real and complex part of each matrix entry), the function giving each component is differentiable.

To motivate this definition, when we studied groups, one important example of a subgroup was *cyclic subgroups*, the subgroups generated by one element. We can think of cyclic subgroups as homomorphisms $\mathbb{Z} \to G$, where the homomorphism maps 1 to the generator of the subgroup. In some sense, this construction makes sense because $\mathbb{Z}$ is the simplest (nontrivial) example of a group — it has one generator and no relations.

Here, the groups we're studying also have some sort of topological structure. The simplest such group is $\mathbb{R}$ — it's "one-dimensional" and has no extra relations. So in this situation we can consider homomorphisms from $\mathbb{R}$ — and we require these homomorphisms to be differentiable so that they play well with the topological structures on both sides.

Example 6.13

When studying SU_2, we saw that for any $x \in \mathbb{E}$, the map $\theta \mapsto \cos\theta \cdot I + \sin\theta \cdot x$ is a homomorphism $\mathbb{R} \to \text{SU}_2$. Its kernel is $2\pi\mathbb{Z}$, and its image is Long_x.

§6.3.1 Matrix Exponentials

First we'll start by trying to find an example of a one-parameter group.

Example 6.14

When $n = 1$, a one-parameter group is a differentiable homomorphism $\varphi\colon \mathbb{R} \to \mathbb{C}^\times$. We can take $\varphi(t) = e^{\alpha t}$ for any $\alpha \in \mathbb{C}$ — this is differentiable, and we have

$$\varphi(s+t) = e^{\alpha(s+t)} = e^{\alpha s}e^{\alpha t} = \varphi(s)\varphi(t).$$

Question 6.15. Is there a version of this construction that works for general n?

The answer is yes — we can define e^A for a *square matrix* A as well. Of course the usual definition of e^x doesn't make sense for matrices, but we also have the power series

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots.$$

This is a very nice power series — it converges everywhere — so we could take it as the *definition* of e^x. And this definition *does* generalize well to matrices:

Definition 6.16. For a $n \times n$ matrix A, its exponential e^A is defined as the $n \times n$ matrix

$$e^A = I + A + \frac{A^2}{2} + \frac{A^3}{3!} + \cdots.$$

Each entry $A^k/k!$ is a $n \times n$ matrix, and it's possible to show that for each of the n^2 matrices, the sum we get is convergent (this can be made more precise by placing a metric on the space of $n \times n$ matrices), so this gives a well-defined $n \times n$ matrix.

Example 6.17

Find e^A and e^B for

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Solution. For the first matrix, we have $A^n = A$ for all $n \geq 1$, so we get

$$e^A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \sum_{n\geq 1} \frac{1}{n!}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix}.$$

For the second, we have $B^2 = 0$, so then

$$e^B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

□

This definition gives us a few useful properties of the matrix exponential:

Proposition 6.18

For any $n \times n$ matrix A and invertible matrix P, we have

$$P^{-1}e^AP = e^{P^{-1}AP}.$$

Proof. This essentially follows considering the power series term-by-term — we have

$$P^{-1}A^kP = (P^{-1}AP)^k$$

for each k. This means the power series of $P^{-1}e^AP$ and $e^{P^{-1}AP}$ are equal if we truncate both at the first k terms, so they're equal in the limit as well. □

This is quite useful — it means it's easy to calculate the exponential of any diagonalizable matrix.

Example 6.19

Find e^A for

$$A = \begin{bmatrix} 0 & 2\pi \\ -2\pi & 0 \end{bmatrix}.$$

Solution. The eigenvalues of A are $\pm 2\pi i$, so we can find some P such that

$$B = PAP^{-1} = \begin{bmatrix} 2\pi i & 0 \\ 0 & -2\pi i \end{bmatrix}.$$

Then we have

$$Pe^AP^{-1} = e^B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

So e^A is conjugate to the identity, which means it must *be* the identity. □

Here we didn't have to calculate what P was, since e^B turned out to be very nice — but in general, it's possible to recover e^A from e^B by calculating P (which is the matrix corresponding to the new eigenbasis).

Proposition 6.20

If v is an eigenvector of A with eigenvalue λ, then v is also an eigenvector of e^A with eigenvalue e^λ.

Proof. We can use a similar argument — we have

$$e^Av = \lim_{k\to\infty}\sum_{\ell=0}^{k}\frac{A^\ell}{\ell!}v = \lim_{k\to\infty}\sum_{\ell=0}^{k}\frac{\lambda^\ell}{\ell!}v = e^\lambda v.$$ □

Finally, the following result will be useful in our analysis of one-parameter groups (since it lets us make use of differentiability):

Proposition 6.21

We have $\frac{d}{dt}e^{tA} = Ae^{tA}$.

Proof. We can again calculate term-by-term using the series — we have

$$\frac{d}{dt}e^{tA} = \frac{d}{dt}\left(I + tA + \frac{t^2}{2}A^2 + \cdots\right),$$

and because of uniform convergence, we can take the term-by-term derivative to get

$$\frac{d}{dt}e^{tA} = 0 + A + tA^2 + \frac{t^2}{2}A^3 + \cdots = Ae^{tA}.$$ □

§6.3.2 Characterization of One-Parameter Groups

Of course, the reason we *started* thinking about matrix exponentials is because we wanted to find a one-parameter group of $\mathrm{GL}_n(\mathbb{C})$, similarly to how $x \mapsto e^x$ was a one-parameter group of $\mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^\times$. The property we need for this to work is the following:

Proposition 6.22

If A and B commute, then $e^{A+B} = e^A e^B$. In particular, we have $e^{(s+t)A} = e^{sA}e^{tA}$.

Proof. By definition, we have

$$e^{A+B} = \sum_{n\geq 0} \frac{(A+B)^n}{n!}.$$

But now we can expand out $(A+B)^n$ by the Binomial Theorem — since A and B commute, in each monomial we can move all the A's to the left, so $(A+B)^n = A^n + \binom{n}{1}A^{n-1}B + \cdots + B^n$, and therefore

$$e^{A+B} = \sum_{n\geq 0} \frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}A^kB^{n-k} = \sum_{k\geq 0}\sum_{\ell\geq 0}\frac{1}{k!}\cdot\frac{1}{\ell!}\cdot A^kB^\ell = \left(\sum_{k\geq 0}\frac{A^k}{k!}\right)\left(\sum_{\ell\geq 0}\frac{B^\ell}{\ell!}\right) = e^Ae^B. \qquad \square$$

In particular, we have $e^A \cdot e^{-A} = e^0 = I$, so then $e^A \in \mathrm{GL}_n(\mathbb{C})$ for *all* matrices $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$.

This proposition immediately implies that the homomorphism $t \mapsto e^{tA}$ is a one-parameter group. But we can also ask if the converse is true:

> **Question 6.23.** Is *every* one-parameter group of the form $t \mapsto e^{tA}$ for some A?

It turns out the answer is yes!

Proposition 6.24

Every one-parameter group in $\mathrm{GL}_n(\mathbb{C})$ is of the form $\varphi\colon t \mapsto e^{tA}$ for a unique matrix A.

So the constraints in the definition of a one-parameter group are a lot stronger than they might seem. This also means there's a bijection between matrices A and one-parameter groups.

Proof. We've already seen that $t \mapsto e^{tA}$ is a one-parameter group — it's differentiable, and it's a homomorphism because $e^{(s+t)A} = e^{sA}\cdot e^{tA}$ for all s and t. So it suffices to prove the other direction.

First we'll prove the uniqueness of A. Suppose φ is the map $t \mapsto e^{tA}$ for some matrix A. Then we have $\varphi'(t) = Ae^{tA}$, so in particular $\varphi'(0) = A$. This means that it's possible to recover A from φ — so given φ, there's *at most* one A for which φ is the map $t \mapsto e^{tA}$. (Note that this result is not as obvious as it may seem, since the map e^A is not injective.)

Now we'll prove the existence of A. Given a one-parameter group φ, set $A = \varphi'(0)$; we'll then show that we must have $\varphi(t) = e^{tA}$ for all t.

The main idea is to obtain a differential equation — we have $\varphi(s+t) = \varphi(s)\varphi(t)$ for all s and t, and taking the derivative with respect to s (while holding t constant) gives

$$\varphi'(s+t) = \varphi'(s)\varphi(t).$$

Now plugging in $s = 0$ gives

$$\varphi'(t) = \varphi'(0)\varphi(t) = A\varphi(t)$$

for all t. So then we have an ordinary differential equation for $\varphi(t)$, and since φ is a homomorphism, we also have the initial condition $\varphi(0) = I$. But it's a fact that the solution to a given first-order differential equation is uniquely determined by one point, so there's only one solution to $\varphi'(t) = A\varphi(t)$ with $\varphi(0) = I$. We already know that e^{tA} is a solution, so this means $\varphi(t) = e^{tA}$. □

The same argument works for $\mathrm{GL}_n(\mathbb{R})$ as well.

§6.4 Lie Algebras

§6.4.1 One-Parameter Subgroups

Given a group G which is a subgroup of $\mathrm{GL}_n(\mathbb{R})$ or $\mathrm{GL}_n(\mathbb{C})$, we can think about the one-parameter groups in G — meaning one-parameter groups $\varphi\colon \mathbb{R} \to \mathrm{GL}_n(\mathbb{C})$ or $\mathrm{GL}_n(\mathbb{R})$ such that $\varphi(t) \in G$ for all t. (This is somewhat similar to the notion of a cyclic subgroup from earlier.)

Example 6.25

The longitudes of $\mathrm{SU}_2 \leq \mathrm{GL}_n(\mathbb{C})$ are one-parameter groups in SU_2, with the homomorphism $\mathbb{R} \to \mathrm{Long}_x$ given by $\theta \mapsto \cos\theta \cdot I + \sin\theta \cdot x$.

Question 6.26. Given a group G, how can we describe all one-parameter groups in G?

Since one-parameter groups are exactly homomorphisms $t \mapsto e^{tA}$, this is equivalent to describing the matrices A for which e^{tA} is in G for all A.

We'll answer this question in a few examples.

Example 6.27

When $G \leq \mathrm{GL}_n(\mathbb{R})$ or $\mathrm{GL}_n(\mathbb{C})$ is the group of diagonal matrices

$$\left\{ \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \;\middle|\; \lambda_i \neq 0 \right\},$$

the one-parameter groups in G are precisely e^{tA} for diagonal matrices A.

Proof. First, any diagonal matrix A does define a one-parameter group in G — if A is diagonal then so is tA for each t, and therefore so is e^{tA}.

On the other hand, suppose we have a one-parameter group $\varphi(t)$, with diagonal entries $\lambda_1(t), \ldots, \lambda_n(t)$. Then by differentiating, we get that $A = \varphi'(0)$ is also diagonal, with entries $\lambda_1'(0), \ldots, \lambda_n'(0)$. So A must be diagonal. □

Remark 6.28. Note that there exist many matrices A for which A is not diagonal but e^{tA} is. So it's important that we know e^{tA} is diagonal for *all* t, since this allows us to differentiate.

Example 6.29

When $G \leq \mathrm{GL}_n(\mathbb{R})$ or $\mathrm{GL}_n(\mathbb{C})$ is the group of upper triangular matrices

$$\left\{ \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \;\middle|\; \lambda_i \neq 0 \right\},$$

the one-parameter groups in G are precisely e^{tA} for upper triangular matrices A.

Proof. This is quite similar to the previous example. First, if A is upper triangular, then A^n is upper triangular for all n, so e^{tA} is upper triangular as well (since it's the sum of upper triangular matrices).

Meanwhile, if $\varphi(t)$ is upper triangular for all t, then so is $\varphi'(t)$, and therefore so is $\varphi'(0) = A$. □

Example 6.30

When $G \leq \mathrm{GL}_n(\mathbb{R})$ or $\mathrm{GL}_n(\mathbb{C})$ is the group of upper triangular matrices with diagonal consisting entirely of 1's, meaning

$$\left\{ \begin{bmatrix} 1 & * & \cdots & * \\ 0 & 1 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \right\},$$

the one-parameter groups in G are precisely e^{tA} for upper triangular matrices A with diagonal consisting entirely of 0's.

Proof. First, to show that any A must be of this form,we can again differentiate at 0 (since the entries of 0 and 1 are all constants, they correspond to entries of 0 in $\varphi'(0) = A$). Meanwhile, if A is of this form, then A^n is upper triangular for all n, and its diagonal entries are all 0's for $n \geq 1$, so then $e^{tA} = I + (tA) + \cdots$ is of the described form. □

Now we'll look at a few harder examples.

Example 6.31

The one-parameter groups in U_n are precisely e^{tA} where A is *skew Hermitian*, meaning that $A^* = -A$.

Recall that $\mathrm{U}_n \leq \mathrm{GL}_n(\mathbb{C})$ is the group of unitary matrices M, matrices for which $M^* = M^{-1}$.

Proof. Exponentiation and taking adjoints behave well with each other — we have

$$(e^A)^* = \left(I + A + \frac{A^2}{2} + \cdots\right)^* = I + A^* + \frac{(A^*)^2}{2} + \cdots = e^{A^*}.$$

Now to show that all such A define a valid one-parameter subgroup, we have $(e^{tA})^* = e^{tA^*}$ since t is real. So if $A^* = -A$, then $e^{tA^*} = e^{-tA} = (e^{tA})^{-1}$, which means e^{tA} is unitary.

On the other hand, if we have $(e^{tA})^* = (e^{tA})^{-1}$ for all t, then we can rewrite this as $e^{tA^*} = e^{-tA}$ for all t. Then taking the derivative at 0 gives $A^* = -A$. □

Example 6.32

The one-parameter groups in O_n are precisely e^{tA} where A is *skew symmetric*, meaning that $A^\intercal = -A$.

The argument here is identical to the previous example — in fact, we can prove this statement *directly* from the previous example, using the fact that $\mathrm{O}_n = \mathrm{U}_n \cap \mathrm{GL}_n(\mathbb{R})$.

Example 6.33

What are the one-parameter subgroups of $\mathrm{SL}_n(\mathbb{C})$?

It turns out that there is a very clean — and surprising — answer!

Lemma 6.34

For any $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$, we have $\det e^A = e^{\operatorname{tr} A}$.

This statement may be unexpected, but it's possible to guess by thinking about diagonal matrices. In fact, that's essentially how we'll prove it as well — it would be hopeless to attempt to prove it directly from definitions, since the determinant doesn't behave well with sums. But we can take advantage of the fact that all the operations involved here behave well with conjugation.

Proof. We know that if A and B are conjugate to each other, then $\operatorname{tr} A = \operatorname{tr} B$, and e^A and e^B are conjugate to each other as well (by Proposition 6.18), so $\det e^A = \det e^B$ as well. So if the statement is true for some matrix A, then it's also true for all matrices conjugate to A.

So then we can assume that A is in Jordan normal form, and is therefore upper triangular, so

$$\begin{bmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

But then as we've seen earlier, e^A is of the form

$$\begin{bmatrix} e^{\lambda_1} & * & * & \cdots & * \\ 0 & e^{\lambda_2} & * & \cdots & * \\ 0 & 0 & e^{\lambda_3} & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{\lambda_n} \end{bmatrix}.$$

So then we have

$$\det e^A = e^{\lambda_1} e^{\lambda_2} \cdots e^{\lambda_n} = e^{\operatorname{tr} A}. \qquad \square$$

Now with this, we can answer our question about $\mathrm{SL}_n(\mathbb{C})$:

Solution to Example 6.33. The matrix A defines a one-parameter group in $\mathrm{SL}_n(\mathbb{C})$ if and only if $\det e^{tA} = 1$ for all $t \in \mathbb{R}$, and by the above lemma this occurs exactly when $e^{\operatorname{tr} tA} = 1$, or equivalently when $\operatorname{tr} tA \in 2\pi i\mathbb{Z}$. But $2\pi i\mathbb{Z}$ is a discrete group and t can be *any* real number, so this only happens when $\operatorname{tr} A = 0$. □

Example 6.35
What are the one-parameter subgroups of SU_2?

Solution. Combining Examples 6.31 and 6.33, the one-parameter groups in SU_n are exactly those defined by matrices A such that $A^* = -A$ and $\operatorname{tr} A = 0$. In the case of SU_2, this means A must be of the form

$$A = \begin{bmatrix} ix_1 & x_2 + ix_3 \\ -x_2 + ix_3 & -ix_1 \end{bmatrix} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}.$$

So then A is a point in $\mathbb{H}$ which can be written as $c\vec{v}$ for some v in the equator of SU_2. Now plugging into the definition of the exponential, we get that

$$e^{tA} = I \cdot \cos tc + v \cdot \sin tc.$$

So if $c = 0$ then this gives just the identity matrix, and otherwise it gives the longitude Long_v. □

Question 6.36. In general, what properties do the matrices A have?

In all these examples, we can see that the set of matrices A is actually a *vector space*! This is surprising, since in *general*, exponentiation doesn't behave well with addition (it only does when the matrices commute).

§6.4.2 Tangent Vectors

To build on this, given a group $G \le \mathrm{GL}_n(\mathbb{R})$ (we can do the same for $\mathrm{GL}_n(\mathbb{C})$ as well), we can consider the set of vectors which are *tangent* to G at the identity matrix. There are a few different ways to define tangent vectors at the identity:

(1) A tangent vector is a $n \times n$ matrix A such that $e^{tA} \in G$ for all $t \in \mathbb{R}$.

(2) Given any differentiable path $f\colon (-\varepsilon, \varepsilon) \to \mathrm{GL}_n(\mathbb{R})$ such that f always lies inside G and $f(0) = I$, the matrix $A = f'(0)$ is a tangent vector.

(3) If G is defined by a bunch of *polynomial* constraints on the entries of its matrices, then there's a more algebraic way of thinking about tangent vectors: work in $\mathbb{R}[\varepsilon]$, where $\varepsilon^2 = 0$. Then we have a more algebraic way of thinking about tangent vectors — if f is a polynomial, then $f(x + \varepsilon) = f(x) + f'(x)\varepsilon$. (The intuition here is that $\varepsilon^2 = 0$ allows us to ignore all higher-degree terms in a power series expansion.) So then we can consider the system of polynomial equations used to define G, and a tangent vector is a $n \times n$ matrix A such that $I + \varepsilon A$ satisfies the same polynomial constraints in $\mathbb{R}[\varepsilon]$.

These definitions are all equivalent (we won't prove this). They have different benefits — the first gives a *bijection* between one-parameter groups and tangent vectors. The second definition has a lot of redundancy, but it makes it easier to see why the set of tangent vectors is always a *vector space*. The third definition is somewhat less general, but many of the groups we've seen so far *are* defined by polynomial constraints. This definition is useful because it makes sense even if we're *not* working over $\mathbb{R}$ or $\mathbb{C}$ — we can talk about tangent vectors to subgroups of $\mathrm{GL}_n(\mathbb{F}_p)$, for instance.

Definition 6.37. The set of tangent vectors to G at the identity is denoted $\mathrm{Lie}(G)$.

Example 6.38

We'll look at the three definitions in the case of $\mathrm{O}_n \leq \mathrm{GL}_n(\mathbb{R})$. We've already seen that using the first definition, $\mathrm{Lie}(\mathrm{O}_n)$ is the set of matrices A for which $A^\intercal = -A$.

Using the second definition, suppose we have a differentiable function $f\colon (-\varepsilon, \varepsilon) \to \mathrm{O}_n$ with $f(0) = I$. Then we have $f(t)^\intercal \cdot f(t) = I$ for all t. Taking the derivative gives

$$f'(t)^\intercal f(t) + f(t)^\intercal f'(t) = 0.$$

Setting $t = 0$ gives $A^\intercal A + IA = 0$, so $A^\intercal = -A$.

Using the third definition, the constraint $M^\intercal M = I$ can be described as a bunch of polynomial constraints on the entries of M. So then tangent vectors are precisely matrices A such that

$$(I + \varepsilon A)^\intercal (I + \varepsilon A) = I.$$

Expanding and using the fact that $\varepsilon^2 = 0$ gives $I + \varepsilon A^\intercal + \varepsilon A + 0 = I$, which means $A^\intercal = -A$.

In this language, our observation that the sets of matrices A are always vector spaces translates to the following statement:

Proposition 6.39

The set $\mathrm{Lie}(G)$ is a vector subspace of $\mathrm{Mat}_{n\times n}(\mathbb{R})$.

§6.4.3 Manifolds

One useful geometric way to think about tangent vectors is in terms of manifolds.

Definition 6.40. Given $M \subset \mathbb{R}^n$, we say M is a *manifold* of dimension d if for each point $x \in M$, there exists an open subset $V \subset M$ containing x, an open ball $U \subset \mathbb{R}^d$, and a bijection $f\colon U \to V$ which is continuous and differentiable.

This definition essentially states that around each point of M, we can find an open subset of M which "looks like" an open ball in $\mathbb{R}^d$ — so locally (but not globally), M looks like $\mathbb{R}^d$.

Example 6.41

The circle $S^1 \subset \mathbb{R}^2$ is a manifold — we can take any small arc and straighten it out into an interval.

Example 6.42

The union of the x-axis and y-axis is *not* a manifold, since any open subset around the origin looks like $+$ instead of an interval.

In fact, all our examples of subgroups of $\mathrm{GL}_n(\mathbb{R})$ are manifolds.

Example 6.43

The group SU_2 is a manifold — we've seen earlier that we can think of it as $S^3 \subset \mathbb{R}^4$. For any point in the upper hemisphere, we can take the open subset $V = \{x_0 > 0\} \cap S^3$, and map the open ball $U = \{x_1^2 + x_2^2 + x_3^2 < 1\} \subset \mathbb{R}^3$ to it by sending

$$(x_1, x_2 x_3) \mapsto \left(\sqrt{1 - x_1^2 - x_2^2 - x_3^2}, x_1, x_2, x_3\right).$$

This makes the fact that the tangent vectors form a vector space a bit more intuitive, since we can "carry over" the vector space of tangent vectors from an open ball U to a neighborhood of I.

§6.4.4 The Lie Bracket and Lie Algebras

As we've seen, $\mathrm{Lie}(G)$ is a vector space, which means we can add and scale its elements. But it consists of *matrices* — it's a subspace of $\mathrm{Mat}_{n\times n}(\mathbb{R})$ — and we don't just know how to add and scale matrices, we also know how to multiply them. Unfortunately, matrix multiplication doesn't work well here — it's possible that A and B are in $\mathrm{Lie}(G)$ and AB is not. But there is a related construction which *does* work well:

Definition 6.44. Given two matrices A and B, their *Lie bracket*, denoted $[A, B]$, is the matrix $AB - BA$.

Of course, if G is abelian, then $[A, B]$ is just 0 on $\mathrm{Lie}(G)$. So in some way, we can think of the Lie bracket as measuring the failure of G to be abelian.

Theorem 6.45

If A and B are in $\mathrm{Lie}(G)$, then so is $[A, B]$.

We'll first look at a few examples:

Example 6.46

In the case of O_n, we have $\mathrm{Lie}(\mathrm{O}_n) = \{A \mid A^\intercal = -A\}$. Now if A and B are both in $\mathrm{Lie}(\mathrm{O}_n)$, then

$$[A, B]^\intercal = B^\intercal A^\intercal - A^\intercal B^\intercal = BA - AB = -[A, B],$$

so $[A, B] \in \mathrm{Lie}(\mathrm{O}_n)$ as well.

Example 6.47

In the case of $\mathrm{SL}_n(\mathbb{R})$, we have $\mathrm{Lie}(\mathrm{SL}_n(\mathbb{R})) = \{A \mid \operatorname{tr} A = 0\}$. But it's true in general that $\operatorname{tr} AB = \operatorname{tr} BA$, so then

$$\operatorname{tr}\,[A, B] = \operatorname{tr} AB - \operatorname{tr} BA = 0,$$

and therefore $\operatorname{tr}\,[A, B] \in \mathrm{Lie}(\mathrm{SL}_n(\mathbb{R}))$.

Sketch of Proof of Theorem 6.45. Suppose we have two matrices $A, B \in \mathrm{Lie}(G)$. Then e^{tA} and e^{tB} are in G for all $t \in \mathbb{R}$, so the matrix

$$e^{tA}e^{sB}e^{-tA}e^{-sB}$$

is also in G. Now we can expand this out using the series definition, to get

$$\left(I + tA + \frac{t^2A^2}{2} + \cdots\right)\left(I + sB + \frac{s^2B^2}{2} + \cdots\right)\left(I - tA + \frac{t^2A^2}{2} - \cdots\right)\left(I - sB + \frac{s^2B^2}{2} - \cdots\right).$$

All the linear terms cancel out, and most of the quadratic terms cancel out as well — we end up with $I + st[A, B] + \cdots$ where the remaining terms have degree at least 3. So since this matrix is in G for all s and t, we can deduce that $[A, B]$ is also a tangent vector to G at I, and is therefore in $\mathrm{Lie}(G)$. □

The Lie bracket satisfies a few important properties: we have $[A, B] = -[B, A]$, and the *Jacobi identity*

$$[[A, B], C] + [[B, C], A] + [[C, A], B] = 0.$$

(This can be proven by expanding out everything, but it also has meaning in terms of the group structure.) So we can define a new algebraic object that comes up naturally from this setting:

> **Definition 6.48.** A *Lie algebra* is a vector space V with a bilinear pairing $[\cdot, \cdot]\colon V \times V \to V$ (called its Lie bracket) which satisfies $[A, B] = -[B, A]$ and the Jacobi identity.

So far, we've only focused on subgroups of $\mathrm{GL}_n(\mathbb{R})$. But we could have performed the same construction for *any* group which is also a manifold — such groups are called *Lie groups*.

It turns out that the Lie algebra of a group carries a lot of information:

> **Theorem 6.49**
>
> Given a finite-dimensional Lie algebra V over $\mathbb{R}$¡ there exists a Lie group G for which $\mathrm{Lie}(G) = V$. Furthermore, if we require that G is simply connected, then there is a *unique* such group G.

> **Remark 6.50.** The condition that G is simply connected is necessary for uniqueness — as shown in the homework, SU_2 and SO_3 have the same Lie algebra. But they "differ by a finite amount" — we've seen that SO_3 is isomorphic to $\mathrm{SU}_2/\{\pm I\}$ — and a similar statement is true in general.

This can be used to understand Lie groups by first understanding Lie algebras (which is often an easier problem).

§6.5 Simple Linear Groups

Recall that a group is *simple* if its only normal subgroups are the trivial group and the entire group. Simple groups are important because in some sense, they're "building blocks" for more complicated groups — since if we have a group which *isn't* simple, we can analyze it by looking at a normal subgroup and its quotient group.

> **Question 6.51.** Are any of our examples of linear groups simple?

We'll focus on two examples — SU_2 and $\mathrm{SL}_2(\mathbb{C})$.

§6.5.1 Normal Subgroups of SU_2

> **Question 6.52.** Is SU_2 simple?

Of course, the answer is no — the center of SU_2 is $\{\pm I\}$, and the center is always a normal subgroup. But it turns out this is essentially the *only* thing that happens.

> **Theorem 6.53**
>
> If N is a normal subgroup of SU_2, then N is either I, $\{\pm I\}$, or SU_2.

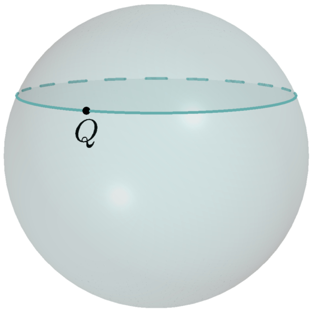
Corollary 6.54

The quotient $\mathrm{SU}_2/\{\pm I\} \cong \mathrm{SO}_3$ is simple.

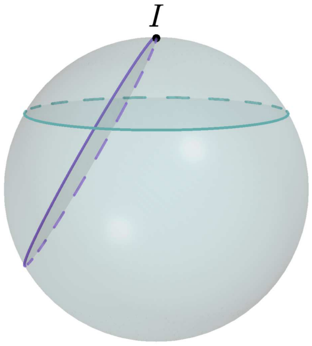
Proof. We can use the correspondence theorem — we have a surjective homomorphism $\varphi\colon \mathrm{SU}_2 \to \mathrm{SO}_3$ (given by mapping each element of SU_2 to the linear operator defined by conjugation on the 3-dimensional space with $x_0 = 0$), whose kernel is $\{\pm I\}$. Then for any normal subgroup of SO_3, its pre-image is a normal subgroup of SU_2 containing $\{\pm I\}$, and the only two such subgroups are $\{\pm I\}$ and SU_2. So then the only normal subgroups of SO_3 are their images $\{I\}$ and SO_3. □

Proof of Theorem 6.53. We'll use our geometric intuition of what SU_2 looks like — it's a 3-sphere, where the latitudes are conjugacy classes and the longitudes are subgroups.

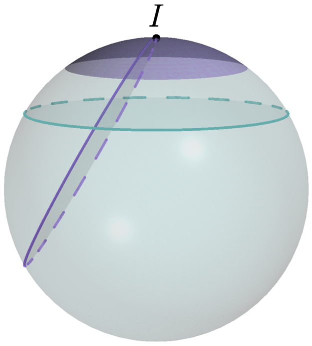
Let N be a normal subgroup of SU_2, and suppose N contains a matrix Q not equal to $\pm I$ — our goal is to show that then $N = \mathrm{SU}_2$. First, since N is normal, it must also contain all elements conjugate to Q, so it contains the entire latitude Lat_c (where $\operatorname{tr} Q = 2c$).



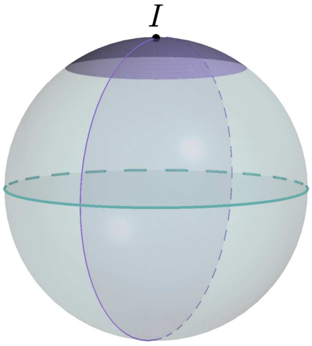
Now we can translate this latitude to pass through the identity — consider $Q^{-1}\,\mathrm{Lat}_c$, which is a (tilted) 2-sphere passing through the north pole. This 2-sphere must be contained in N as well.



Now we can take a nontrivial path $f(t)$ starting at I and staying in $Q^{-1}\,\mathrm{Lat}_c$. This path must be contained in N, and it must contain *some* matrix of every trace in some interval $(2-\delta, 2]$, for some δ. But since N is normal and the conjugacy classes are precisely the latitudes, then N must contain *all* such matrices!



Now we're almost done, and to finish, we can consider the *longitudes*.



For each $v \in \mathbb{E}$, we can look at the longitude through v, which is a subgroup of SU_2. Since every point in SU_2 is in *some* longitude, it suffices to show that Long_v is contained in N. But Long_v consists of elements $\rho_\theta = \cos\theta \cdot I + \sin\theta \cdot \vec{v}$, and we know ρ_θ is in N for all θ in some nontrivial interval $(-\varepsilon, \varepsilon)$. But then for *any* angle φ, we can find some positive integer m such that $|\varphi/m| < \varepsilon$, and then since $\rho_{\varphi/m}$ is in N, so is its mth power ρ_φ.

So for all $v \in \mathbb{E}$, the longitude Long_v is contained in N; since every point in SU_2 is in some longitude, this means $N = \mathrm{SU}_2$. □

§6.5.2 Normal Subgroups of SL_2

Question 6.55. Is $\mathrm{SL}_2(\mathbb{C})$ simple?

Of course, the answer is again no — its center is $\{\pm I\}$, so the most we could ask for is that the quotient $\mathrm{SL}_2(\mathbb{C})/\{\pm I\}$ is simple. It turns out that this *is* true, and in fact, it's true for almost *any* field, not just $\mathbb{C}$!

Theorem 6.56

For any field F with $|F| \geq 4$, the quotient group $\mathrm{SL}_2(F)/\{\pm I\}$ is simple.

This quotient group has a name — it's called $\mathrm{PSL}_2(F)$.

Since we're no longer in a geometric setting — F can even be a finite field — the proof won't be geometric like in the case of SU_2. Instead, we'll look at conjugacy classes and attempt to use them to generate all of the group.

Remark 6.57. The theorem is false for $\mathbb{F}_2$ and $\mathbb{F}_3$ — this is somewhat similar to how A_n is simple for all $n \geq 5$, but the smaller groups A_n aren't simple.

Proof. We'll assume that $|F| > 5$ — there's only two cases this doesn't cover, and they can be checked by hand. Similarly to the case of SU_2, it suffices to prove that the only normal subgroups of $\mathrm{SL}_2(F)$ are $\{I\}$, $\{\pm I\}$, and $\mathrm{SL}_2(F)$ itself.

Claim — Given any $a \in F$, the equation $x^2 = a$ has at most two solutions.

Proof. If $x^2 = y^2$, then $(x+y)(x-y) = 0$. But since F is a field, this implies $x+y$ or $x-y$ must be 0. This means if we have *one* solution to $x^2 = a$, there's at most one other solution (which is $-x$). ■

Claim — If $|F| > 5$, then there exists some $r \in F$ such that $r^2 \notin \{0, \pm 1\}$.

Proof. There are one square root of 0, two square roots of 1, and at most two square roots of -1; since $|F| > 5$, this means there must be an element r which is not a square root of any of these numbers. ■

Now fix some r with $r^2 \notin \{0, \pm 1\}$. Suppose we have a normal subgroup N which contains some element other than $\pm I$ — so we want to show N is the entire group $\mathrm{SL}_2(F)$.

Claim — There exists some $B \in N$ with distinct eigenvalues.

Proof. Take some $A \in N$ with $A \neq \pm I$. Then A cannot be a scalar matrix, so there is some vector $v_1 \in F^2$ which is not an eigenvector of A. Let $v_2 = Av_1$, so then v_1 and v_2 form a basis for F^2 (since they're not linearly dependent).

Now define $P \in \mathrm{GL}_2(F)$ with the property that $Pv_1 = rv_1$ and $Pv_2 = r^{-1}v_2$. Then the eigenvalues of P are r and r^{-1}, so $\det P = 1$ and therefore $P \in \mathrm{SL}_2(F)$.

Now we can take $B = APA^{-1}P^{-1}$. This must be in N — since A is in N, so is A^{-1}, and then since N is normal, so is its conjugate $PA^{-1}P^{-1}$. But we have

$$Bv_2 = APA^{-1}P^{-1}v_2 = APA^{-1}rv_2 = APrv_1 = Ar^2v_1 = r^2v_1,$$

so then r^2 is an eigenvalue of B, and since $\det B = 1$, its other eigenvalue must be r^{-2}. Since $r^2 \neq \pm 1$, we have $r^2 \neq r^{-2}$, so the eigenvalues of B are indeed distinct. ■

Now let $s = r^2$, so B has eigenvalues s and s^{-1}.

Claim — The matrices in $\mathrm{SL}_2(F)$ with eigenvalues s and s^{-1} form a conjugacy class.

Proof. Take any such Q. Then Q has distinct eigenvalues, so it is diagonalizable, which means we can find $L \in \mathrm{GL}_2(F)$ for which

$$LQL^{-1} = \begin{bmatrix} s & 0 \\ 0 & s^{-1} \end{bmatrix}.$$

So these matrices are conjugate to each other in $\mathrm{GL}_2(F)$, but we can in fact show they're conjugate in $\mathrm{SL}_2(F)$ as well by choosing L to have determinant 1 — given any L with $\det L = \delta$, we can take

$$\widetilde{L} = \begin{bmatrix} \delta^{-1} & 0 \\ 0 & 1 \end{bmatrix} L,$$

which is in $\mathrm{SL}_2(F)$ and has the same property. ■

So then since N contains *one* matrix B with eigenvalues s and s^{-1}, it must contain *all* of them. Finally, we can show that these matrices generate $\text{SL}_2(F)$ — for example, it's possible to write down explicit formulas producing the elementary matrices

$$\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 \\ x & 1 \end{bmatrix}$$

for all x, and these matrices generate $\text{SL}_2(F)$ (by using row reduction, for instance). □

Although these two proofs had quite different settings, they had the same general idea — we find an element in N, conjugate it to find a whole bunch of elements in N, and use these elements to generate the entire group.

Remark 6.58. These examples of simple linear groups actually generalize to higher dimensions. In fact, for linear groups defined by *polynomial* constraints (for example, the determinant is a polynomial in the entries, but complex conjugation isn't), there's actually a full classification of which ones are simple. For example, SO_n and SL_n mod their centers work. The proof involves Lie algebras — you first understand what the Lie algebra of a simple group looks like, and use that to characterize the groups.

It's also possible to use this classification to produce *finite* simple groups, by taking F to be a finite field instead of $\mathbb{R}$ or $\mathbb{C}$.

§7 Hilbert's Third Problem

§7.1 Polygons in the Plane

Definition 7.1. Two polygons P and Q are *scissors-congruent* if using finitely many cuts, we can divide each of P and Q into the same collection of polygons $R_1, \dots, R_n$.

In other words, we can cut P up into pieces and rearrange these pieces to form Q. If P and Q are scissors-congruent, we denote this by $P \sim Q$.

Question 7.2. Given two polygons P and Q, when are they scissors-congruent?

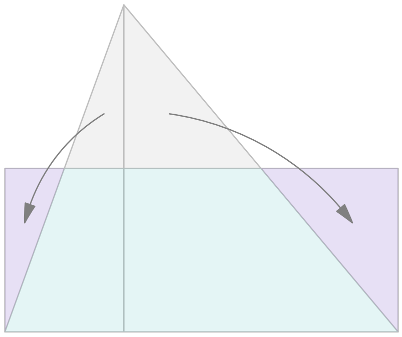
Of course, if $P \sim Q$, then P and Q must have the same area. It turns out this is the *only* obstruction:

Theorem 7.3

If P and Q have the same area, then $P \sim Q$.

Proof Sketch. We'll show that if P has area a, then P is scissors-congruent to the rectangle of dimensions $1 \times a$; then it follows that P and Q are scissors-congruent to the same shape, and therefore to each other.

First, we can cut P into triangles $T_1, \dots, T_n$. Then each triangle is scissors-congruent to *some* rectangle:



Then it's possible to show that any rectangle with area c is scissors-congruent to a rectangle with dimensions $1 \times c$ (this part is finicky and involves a lot of cases, so we'll skip it). Then we can concatenate all our height-1 rectangles to get that P is scissors-congruent to a $1 \times a$ rectangle. □

§7.2 Hilbert's Third Problem

Question 7.4. What happens in three dimensions?

Now instead of polygons, we work with *polytopes* (or polyhedra) — which have finitely many vertices, edges, and faces. The definition of scissors-congruence is the same — two polytopes P and Q are scissors-congruent if we can use finitely many straight cuts to decompose P and Q into the same polytope pieces.

Of course, we have the same obvious constraint on scissors-congruence as in the two-dimensional case — if $P \sim Q$, then they must have the same volume. The question we'll study today is the following:

Question 7.5 (Hilbert's Third Problem). If two polytopes have the same volume, are they necessarily scissors-congruent?

As some historical background, in 1900 Hilbert made a list of around twenty problems, which he considered the most important problems in modern mathematics. This was one of the problems on the list, and he expected the answer was no. In fact, this was the first problem to be answered — in 1901, by his student Max Dehn. More precisely, Dehn showed that a cube and a tetrahedron of the same volume are not scissors-congruent.

§7.3 The Tensor Product

At the heart of this problem is a certain algebraic construction.

Definition 7.6. Given two abelian groups G and H, their *tensor product* is the abelian group $G \otimes H$ generated by elements denoted by $g \otimes h$ for $g \in G$ and $h \in H$, which satisfy the relations

$$(g + g') \otimes h = g \otimes h + g' \otimes h$$
$$g \otimes (h + h') = g \otimes h + g \otimes h'.$$

One way to think of the tensor product $G \otimes H$ is as

$$\bigoplus_{g,h} \mathbb{Z}(g \otimes h)/S$$

where S is the subgroup generated by all the elements $(g+g')\otimes h - g\otimes h - g'\otimes h$ and $g\otimes(h+h') - g\otimes h - g\otimes h'$ — we essentially take *all* formal linear combinations of elements $g \otimes h$, and quotient out by all the relations. Intuitively, the elements of $G \otimes H$ are all combinations of the terms $g \otimes h$, which we know how to simplify.

Our definition has a few immediate consequences:

Proposition 7.7

The tensor product has the following properties:

- $0 \otimes g = g \otimes 0 = 0$.
- For any integer a, we have $(ag) \otimes h = a(g \otimes h) = g \otimes (ah)$. (Here ag denotes g added to itself a times.)
- If G is generated by $g_1, \dots, g_r$ and H by $h_1, \dots, h_s$, then $G \otimes H$ is generated by $g_i \otimes h_j$ over all i and j (since we can use the relations repeatedly to reduce any $g \otimes h$ to a sum of such terms). This works even if the set of generators is not finite.

Example 7.8

For any group G, we have $\mathbb{Z} \otimes G \cong G$, with the isomorphism $a \otimes g \mapsto ag$.

Example 7.9

For any group G, we have $\mathbb{Z}^2 \otimes G \cong G \times G$, with the isomorphism $(a, b) \otimes g \mapsto (ag, bg)$.

Note that in these two examples, we could write all elements of $G \otimes H$ in the form $g \otimes h$, but this isn't true in general — we can have elements such as $g_1 \otimes h_1 + g_2 \otimes h_2$ which can't be simplified any further.

Example 7.10

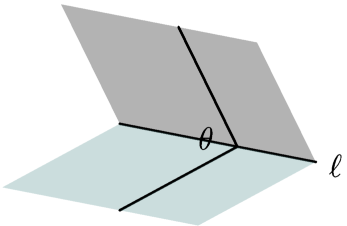
We have $C_2 \otimes C_3 = 0$ — to see this, take any $x \otimes y$. Then we have $3x = x$ (since $x \in C_2$) and $3y = 0$ (since $y \in C_3$), so

$$x \otimes y = 3x \otimes y = x \otimes 3y = x \otimes 0 = 0.$$

So it's possible to tensor together two nontrivial groups and end up with a trivial one — so the tensor product is somewhat subtle.

§7.4 The Dehn Invariant

To answer our question, we want another property of polytopes which is preserved under scissors congruence. Given a polytope, each edge has a length $\ell \in \mathbb{R}$ and a *dihedral angle* $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ (the angle between the two faces that meet at the edge — or more precisely, the angle between the perpendiculars to the edge on those two faces).



Then $\ell \otimes \theta$ defines an element in $\mathbb{R} \otimes \mathbb{R}/2\pi\mathbb{Z}$.

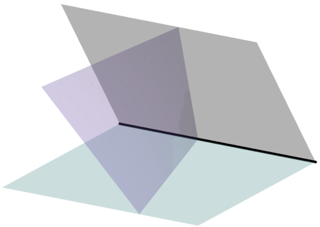
Definition 7.11. The *Dehn invariant* of a polytope P, denoted $d(P)$, is the sum of $\ell_i \otimes \theta_i$ over all edges i in P.

Theorem 7.12

The Dehn invariant is preserved by scissors congruence — if $P \sim Q$ then $d(P) = d(Q)$.

Proof Sketch. It suffices to show that cutting the polytope preserves scissors-congruence. When we cut, there's a few different things that can happen. We won't carefully go through all the cases, but we'll see a few of them to see why this "should" be true.

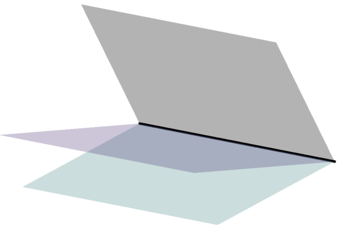
Case 1 (We cut an edge into two pieces, keeping the dihedral angle on both sides the same).



Then we originally have one edge (ℓ, θ), and we end up with two edges (ℓ_1, θ) and (ℓ_2, θ). But $\ell_1 + \ell_2 = \ell$, so

$$\ell_1 \otimes \theta + \ell_2 \otimes \theta = (\ell_1 + \ell_2) \otimes \theta = \ell \otimes \theta.$$

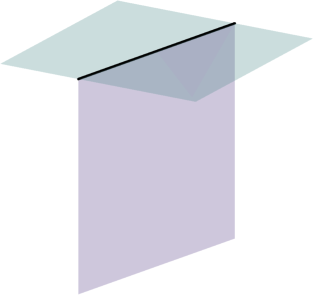
Case 2 (We cut along a plane containing the edge, preserving the length but splitting the dihedral angle).



Then we start with (ℓ, θ) and end up with (ℓ, θ_1) and (ℓ, θ_2), where $\theta_1 + \theta_2 = \theta$, so

$$\ell \otimes \theta_1 + \ell \otimes \theta_2 = \ell \otimes (\theta_1 + \theta_2) = \ell \otimes \theta.$$

Case 3 (We create an edge on the outside of the polytope, by cutting through a face).



Then we started with no edge, and we produced (ℓ, θ_1) and (ℓ, θ_2), where $\theta_1 + \theta_2 = \pi$. Here we have

$$\ell \otimes \theta_1 + \ell \otimes \theta_2 = \ell \otimes \pi = \frac{\ell}{2} \otimes 2\pi = 0.$$

It's possible to cover the remaining cases similarly. □

This now gives us a property other than volume which is invariant under scissors-congruence. But for this to be useful, we need to check that it actually *does* differentiate between polytopes — it happens scarily often that a complicated invariant turns out to just always be 0.

Theorem 7.13

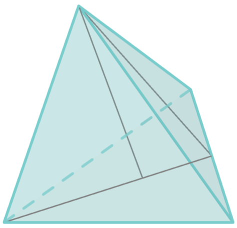
Any cube and regular tetrahedron have different Dehn invariants.

Proof. Call the cube C and tetrahedron T. Then C has 12 edges, each with dihedral angle $\pi/2$. So then

$$d(C) = 12 \cdot \ell \otimes \frac{\pi}{2} = \ell \otimes 6\pi = 0.$$

So any cube has Dehn invariant 0.

On the other hand, T has 6 edges of some length ℓ (not necessarily the same as the edge lengths of the cube) and the same dihedral angle α, so $d(T) = 6 \cdot \ell \otimes \alpha$. To find α, we can drop a few perpendiculars:



Then we can see that $\cos\alpha = \frac{1}{3}$ (since all the triangles have the same height, and the center of a face occurs $\frac{2}{3}$ of the way along its height).

> **Claim** — α is not a rational multiple of π.

Proof. If α were a rational multiple of π, then we would have $\cos n\alpha = 1$ for some positive integer n. But by trig identities, $\cos n\alpha$ is a polynomial in $\cos\alpha$ with leading coefficient 2^{n-1}. So $\cos n\alpha$ must have a power of 3 in its denominator, contradiction. ■

> **Claim** — If α is not a rational multiple of π and ℓ is nonzero, then $\ell \otimes \alpha$ is nonzero.

Proof. We can think of $\mathbb{R}$ as a vector space over $\mathbb{Q}$ (with uncountable dimension). Then π and α are linearly independent, so we can fill them out into a basis for $\mathbb{R}$ — we can write $\mathbb{R} = \mathbb{Q}\alpha + \mathbb{Q}\pi + W$ for a $\mathbb{Q}$-vector space W. Then we can define the linear map $f\colon \mathbb{R} \to \mathbb{Q}$ (as a map between $\mathbb{Q}$-vector spaces) sending $\alpha \mapsto 1$, and every other basis element to 0.

This gives a group homomorphism $\mathbb{R} \otimes \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{R}$ defined by $z \otimes x \mapsto zf(\widetilde{x})$, where $\widetilde{x}$ is x mod 2π (note that $f(\widetilde{x})$ is well-defined because 2π is in the kernel of f). Then $\ell \otimes \alpha$ is mapped to ℓ, which is nonzero; so $\ell \otimes \alpha$ must be nonzero as well. ■

This implies $d(T)$ is nonzero, and therefore $d(C) \neq d(T)$. □

So a cube and tetrahedron cannot be scissors-congruent, even if they have the same volume.

> **Remark 7.14** (Historical Note)**.** This result was proven in 1901 by Dehn. In 1968, Sydler showed the converse — if two polytopes have the same volume *and* Dehn invariant, then they're scissors-congruent. The same result is true in four dimensions as well, but we don't have a characterization in higher dimensions of when two polytopes are scissors-congruent.