PMATH 347 LECTURES

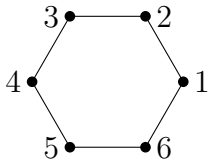
ROSS WILLARD
UNIVERSITY OF WATERLOO
FALL 2014

Contents

1. Sept 8 – Dihedral symmetries

For $n \geq 3$, let C_n denote a *regular n-gon* (embedded in $\mathbb{R}^3$). A *dihedral symmetry* of C_n is any "rigid motion" of $\mathbb{R}^3$ which moves C_n back to itself.

For example, let $n = 6$:



Dihedral symmetries of C_6 include:

- Rotations around the center for C_6 (by multiples of 60°).
- "Flips" (called *reflections*) along an axis – either through two opposite vertices, or through the centers of two opposite sides.
- The "identity" symmetry (which does nothing).

Definition. D_{2n} = the set of all dihedral symmetries of C_n.

Note: In geometry the set is called D_n.

D_{2n} clearly includes:

- n rotations (including the identity symmetry), by multiples of $2\pi/n$ radians.
- n reflections.

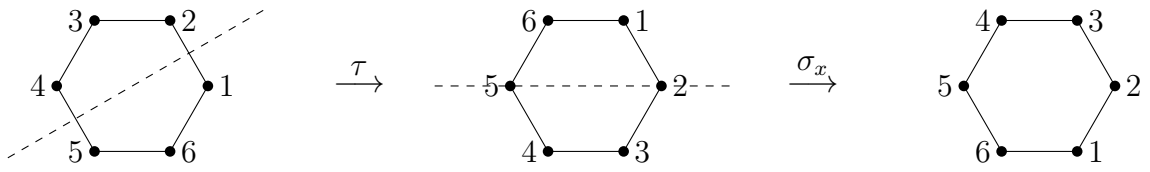
Let's prove that there are no other dihedral symmetries than these.

Lemma. $|D_{2n}| = 2n$.

Proof. We've already seen that $|D_{2n}| \geq 2n$. To prove the opposite inequality, suppose σ is an arbitrary dihedral symmetry of C_n. Then σ must send 1 to some vertex $i \in \{1, 2, \ldots, n\}$. Since σ preserves edges, it must send 2 to $i+1$ or $i-1$. Thus there are only $n \times 2$ possibilities for $(\sigma(1), \sigma(2))$. Since $\sigma(1), \sigma(2)$ determine σ (think about it), we have $|D_{2n}| \leq 2n$. □

We can combine or "multiply" dihedral symmetries, by applying one first and then the other. Convention: $\sigma \cdot \tau$ means "τ first, then σ."

Example. Let σ_x be the reflection through the x-axis, and τ the reflection through the line through the centers of 12 and 45. Then $\sigma_x \cdot \tau$ is



Thus $\sigma_x \cdot \tau$ = rotation **clockwise** by 60°.

Exercise: check that $\tau \cdot \sigma_x$ = rotation **counter-clockwise** by 60°.

We adopt the following convention (slightly different from the text). When n is understood, we let:

- r denote the rotation counter-clockwise by $2\pi/n$ radians.
- s denote reflection through the x-axis.
- 1 denote the identity symmetry.

Thus r^2 $(= r \cdot r)$ is rotation by $4\pi/n$ radians ccw, r^3 is rotation by $6\pi/n$ radians ccw, etc. $r^n = 1$. What is r^{n-1}? We also write this symmetry as r^{-1}.

One can check that:

- sr = reflection through the line through center of $1n$.
- sr^2 = reflection through the line through n.
- sr^3 = reflection through the line through the center of $n-1, n$.
- Etc.

Thus $s, sr, sr^2, \ldots, sr^{n-1}$ enumerate all n reflections. Hence we can write

$$D_{2n} = \{r^i \,:\, 0 \leq i < n\} \cup \{sr^i \,:\, 0 \leq i < n\}.$$

The expressions $1, r, r^2, \ldots, r^{n-1}, s, sr, sr^2, \ldots$ are called the *normal forms* for the symmetries they represent.

In addition:

- $rs = sr^{n-1}$ $(= sr^{-1})$.

This last fact gives us an easy way to multiply two symmetries. E.g.,

$$sr^2 \cdot sr = srrsr = (sr)(rs)r = (sr)(sr^{-1})r = s(rs) = s(sr^{-1}) = r^{-1}.$$

(Note that $s^2 = 1$.) In fact, all that we need to know to calculate products is $\boxed{r^n = s^2 = 1 \text{ and } rs = sr^{-1}}$.

2. Sept 9 – Permutations

Definition. Let X be any non-empty set.

(1) A *permutation* of X is a bijection $\sigma : X \to X$.
(2) S_X is the set of all permutations of X.
(3) If $X = \{1, 2, 3, \ldots, n\}$ then we denote S_X by S_n.

What is $|S_n|$? (Answer: $n!$) Proof: there are n choices for $\sigma(1)$; once chosen, there are $n-1$ choices for $\sigma(2)$; etc. The total number of choices for values is $n \cdot (n-1) \cdots 2 \cdot 1$.

When X is finite we use a special notation to describe permutations. Example: let $\sigma \in S_8$ be the function

x	1	2	3	4	5	6	7	8
$\sigma(x)$	4	8	6	3	5	1	2	7

Start with $x = 1$; where does σ send it? to 4. Then where does σ send 4? to 3. Repeat:

$$1 \mapsto 4 \mapsto 3 \mapsto 6 \mapsto 1.$$

We encode this information as $(1\,4\,3\,6)$. It gives the values of σ at 1, 3, 4 and 6.

We can start with 2: $2 \mapsto 8 \mapsto 7 \mapsto 2$. We record this information as $(2\,8\,7)$.

All that is left is the element 5. $5 \mapsto 5$ and we record this as (5).

We put it all together and write

$$\sigma = (1\,4\,3\,6)(2\,8\,7)(5).$$

This notation completely specifies σ. Note that we could have also written

$$\begin{aligned} \sigma &= (2\,8\,7)(5)(1\,4\,3\,6) \\ &= (3\,6\,1\,4)(5)(7\,2\,8) \end{aligned}$$

etc. There is no uniqueness of notation. What is unique is the individual *cycles*. In this example we say that σ *decomposes* into one 4-cycle, one 3-cycle, and one 1-cycle. Also note that in cycle notation, **the cycles are (pairwise) disjoint**.

Convention: We do not write 1-cycles. Thus in this example, $\sigma = (1\,4\,3\,6)(2\,8\,7)$.

Inverses. Note that if $\sigma \in S_X$, then σ^{-1} exists and is also in S_X. The cycle notation for σ^{-1} is easy: just reverse the cycles of σ. E.g., if

$$\sigma = (1\ 4\ 3\ 6)(2\ 8\ 7)$$

then

$$\sigma^{-1} = (1\ 6\ 3\ 4)(2\ 7\ 8).$$

Composition. Suppose $\sigma, \tau \in S_X$. So they are both functions $X \to X$. So we can compose them to get another permutation $\sigma \circ \tau : X \to X$, which we write as $\sigma\tau$.

For example, suppose σ is as before and let $\tau = (2\ 4\ 8\ 5)(1\ 7)$. In other words, τ is the function

x	1	2	3	4	5	6	7	8
$\tau(x)$	7	4	3	8	2	6	1	5

Let's find $\sigma\tau$. We have

$$\begin{aligned}(\sigma\tau)(1) &= \sigma(\tau(1)) = \sigma(7) = 2\\ (\sigma\tau)(2) &= \sigma(\tau(2)) = \sigma(4) = 3\\ (\sigma\tau)(3) &= \sigma(\tau(3)) = \sigma(3) = 6\end{aligned}$$

and so on. The full table is

x	1	2	3	4	5	6	7	8
$(\sigma\tau)(x)$	2	3	6	7	8	1	4	5

Now let's find the cycle notation for $\sigma\tau$ (from the table). We see that

$$\sigma\tau = (1\ 2\ 3\ 6)(4\ 7)(5\ 8).$$

Actually, we could have found this directly from the cycle notations for σ and τ, as follows.

(1) Write the cycle notation for σ followed by the notation for τ:

$$\sigma\tau = \underbrace{(1\,4\,3\,6)(2\,8\,7)}_{\sigma}\underbrace{(2\,4\,8\,5)(1\,7)}_{\tau}.$$

(Note that this expression is **not** cycle notation.)

(2) Start with 1; reading the cycles from **right to left**, find the first cycle that moves 1 ($1 \mapsto 7$). Continuing to the left, find the first cycle that moves 7 ($7 \mapsto 2$); continue to the left, find the first cycle that moves 2 (there is none). So $\sigma\tau$ sends $1 \mapsto 2$. Continue in this way to find

$$\sigma\tau = (1\ 2\ 3\ 6)(4\ 7)(5\ 8).$$

Let's compute $\tau\sigma$.

$$\tau\sigma = \underbrace{(2\ 4\ 8\ 5)(1\ 7)}_{\tau}\underbrace{(1\,4\,3\,6)(2\,8\,7)}_{\sigma} = (1\ 8)(2\ 5)(3\ 6\ 7\ 4).$$

Note that $\sigma\tau \neq \tau\sigma$. (However $\sigma\tau$ and $\tau\sigma$ do have the same "cycle structure." Is this always true?)

Special notation, terminology.

(1) 1 denotes the identity permutation in S_X (so $1(x) = x$ for all $x \in X$).
(2) The cycle notation for 1 is . (Empty)
(3) Given $\sigma \in S_X$, the **support** of σ is the set

$$\mathrm{supp}(\sigma) = \{x \in X \,:\, \sigma(x) \neq x\}.$$

Equivalently, $\mathrm{supp}(\sigma)$ is the set of elements mentioned in the cycle notation of σ.

(4) σ, τ are **disjoint** if $\mathrm{supp}(\sigma) \cap \mathrm{supp}(\tau) = \varnothing$.

3. Sept 9 – Definition of a group

Definition. Let A be a non-empty set. A *binary operation on A* is a function $*$ whose domain is $A \times A$ (the set of all ordered pairs from A) and which maps into A.

Notational convention: we write $a * b$ for the value of $*$ at (a, b), instead of writing $*(a, b)$.

Definition 3.1. A *group* is an ordered pair $(G, *)$ where

- G is a non-empty set;
- $*$ is a binary operation on G;

which jointly satisfy the following further conditions:

(i) $*$ is *associative*: $(a * b) * c = a * (b * c)$ for all $a, b, c \in G$;
(ii) There exists an *identity* element $e \in G$: $a * e = e * a = a$ for all $a \in G$;
(iii) Every $a \in G$ has a 2-sided *inverse*, i.e., an element $a' \in G$ which satisfies $a * a' = a' * a = e$.

Examples

(1) $(\mathbb{Z}, +)$.
  - $a, b \in \mathbb{Z}$ implies $a + b \in \mathbb{Z}$.
  - $+$ is associative.
  - $0 \in \mathbb{Z}$ satisfies $a + 0 = 0 + a = 0$ for all $a \in \mathbb{Z}$.
  - Every $n \in \mathbb{Z}$ has an inverse $-n \in \mathbb{Z}$, and $n + (-n) = (-n) + n = 0$.

(2) $(D_{2n}, \cdot)$ for each $n \geq 3$. (Called the *dihedral group of order* $2n$.)
  - $\sigma, \tau \in D_{2n}$ implies $\sigma \cdot \tau \in D_{2n}$.
  - $\cdot$ is associative (e.g. because it is composition of functions).
  - 1 (the identity symmetry) satisfies $\sigma \cdot 1 = 1 \cdot \sigma = \sigma$ for all $\sigma \in D_{2n}$.
  - Every $\sigma \in D_{2n}$ has an inverse symmetry $\sigma^{-1} \in D_{2n}$ (doing σ in reverse), and $\sigma \cdot \sigma^{-1} = \sigma^{-1} \cdot \sigma = 1$.

(3) $(S_n, \circ)$ for each $n \geq 2$. (Called the *symmetric group of degree n.*)
  - $\sigma, \tau \in S_n$ implies $\sigma \circ \tau \in S_n$.
  - $\circ$ is associative (because it is composition of functions).
  - 1 (the identity permutation) satisfies $\sigma \circ 1 = 1 \circ \sigma = \sigma$ for all $\sigma \in S_n$.
  - Every $\sigma \in S_{2n}$ has an inverse function $\sigma^{-1} \in S_n$ and $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = 1$.

(4) ($\{$all invertible $n \times n$ matrices over $\mathbb{R}\}$, matrix multiplication). Called $GL_n(\mathbb{R})$.

(5) ($\mathbb{Z}_n$, + mod n).

The groups in (2), (3) and (5) are *finite*, while those in (1) and (4) are *infinite*.

Notation:

- Denote a group $(G, *)$ by G.
- Write ab for $a * b$ (most of the time).
- Denote the identity element of G by 1 (most of the time).

- Denote the inverse a' of an element a by a^{-1} (most of the time).
- The **order** of a group G, denoted $|G|$, is the number of its elements.
- A group $(G, *)$ is *abelian* if $*$ is commutative ($a * b = b * a$ for all $a, b \in G$), and is *non-abelian* otherwise.

Powers

Definition. In any group G, if $a \in G$ then define $a^0 = 1$ and $a^{n+1} = a \cdot a^n$ for $n \geq 0$. Also define $a^{-n} = (a^n)^{-1}$ for $n \geq 1$.

This notation satisfies the usual rules for exponents:

Fact. Let $(G, \cdot)$ be a group, $a \in G$, and $m, n \in \mathbb{Z}$.

(1) $a^1 = a$.
(2) $a^m \cdot a^n = a^{m+n}$.
(3) $(a^m)^n = a^{mn}$.
(4) $a^{-n} = (a^n)^{-1} = (a^{-1})^n$.

The proofs are by induction and case-analysis, depending on whether $m, n > 0$, $= 0$, or < 0, heavily using associativity.

Warning. In general, it is not true that $(ab)^n = a^n b^n$. E.g., $(ab)^2 = abab$, $a^2b^2 = aabb$. We would need $ba = ab$, which is not always true.

Warning. In groups for which the operation is addition, this notation can be alarming. E.g., in the group $(\mathbb{R}, +)$, if $a \in \mathbb{R}$ and $n \geq 2$, then $a^n = \underbrace{a + a + \cdots a}_{n}$, which we prefer to write as na. (Similarly, a^1 is $1a = 1$, a^{-1} is $-a$ which we write as $(-1)a$, and a^0 is $0a = 0$.)

Finite groups exhibit *periodicity* in the following way. Suppose $a \in G$ and consider $a, a^2, a^3, a^4, \ldots,$. If G is finite, then there must be a repeat, say $a^i = a^j$ with $i < j$. Multiply both sides by a^{-1} to get $a^{i-1} = a^{j-1}$. Repeat i times until getting $1 = a^{j-i}$. This proves the existence of $n > 0$ such that $a^n = 1$. Then $a^{n+1} = a^n a^1 = 1 \cdot a = a$, $a^{n+2} = a^2$, etc.

Definition. For a group G and element $a \in G$, the *order* of a (denoted $|a|$ or $\circ(a)$) is the least integer $n > 0$ such that $a^n = 1$, if it exists. If no such n exists, then the order of a is defined to be ∞.

Examples.

- In D_{2n}, the rotation r has order n because $r \neq 1$, $r^2 \neq 1$, $\ldots$, $r^{n-1} \neq 1$ but $r^n = 1$. [Question: what is the order of a reflection in D_{2n}?]
- In $\mathbb{Z}_4$, what is the order of 1? ($1 \neq 0$, $1 + 1 \neq 0$, $1 + 1 + 1 \neq 0$, but $1 + 1 + 1 + 1 = 0$.)

4. Sept 15 – Elementary properties of groups

Proposition 1.2. *Let G be a group and $a, b, u, v \in G$.*

(1) *Left and right cancellation:*
   (a) *If $au = av$, then $u = v$.*
   (b) *If $ub = vb$, then $u = v$.*
(2) *The equations $ax = b$ and $ya = b$ have unique solutions for $x, y \in G$.*

Proof. (1) Assume $au = av$. Then $a^{-1}(au) = a^{-1}(av)$. Thus

$$u = 1u = (a^{-1}a)u = a^{-1}(au) \stackrel{*}{=} a^{-1}(av) = (a^{-1}a)v = 1v = v.$$

The proof of right cancellation is similar.

(2) Let $x := a^{-1}b$. x is one solution to $ax = b$: $a(a^{-1}b) = (aa^{-1})b = 1b = b$. Conversely, suppose u is a solution. Then $au = ax$, so $u = x$. Similarly, $y := ba^{-1}$ is the unique solution to $ya = b$. □

Corollary. *In any group G, the identity element is unique.*

Proof. Suppose d, e are both identity elements of G. Thus $xd = dx = x$ and $xe = ex = x$ for all $x \in G$. In particular, $xd = xe$. Hence $d = e$ by left cancellation. □

Similar tricks let us prove the following.

Proposition 1.1. *Suppose G is a group.*

(1) *Each $a \in G$ has a unique inverse a^{-1}.*
(2) *$(a^{-1})^{-1} = a$ for all $a \in G$.*
(3) *$(ab)^{-1} = (b^{-1})(a^{-1})$ for all $a, b \in G$.*

Proof. (1) Suppose that a has two inverses a' and a^*. That means $aa' = a'a = 1$ and $aa^* = a^*a = 1$. Find a way to use left cancellation to deduce $a' = a^*$.

(2) Let $b = a^{-1}$. We are required to show that $b^{-1} = a$. We know that $bb^{-1} = b^{-1}b = 1$ and $ab = ba = 1$. Use left cancellation.

(3) We'll use a similar trick. Let $c = ab$ and $d = b^{-1}a^{-1}$. We want to show $c^{-1} = d$, so it suffices to show $cc^{-1} = cd$. Well, $cc^{-1} = 1$, while

$$cd = (ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = a1a^{-1} = (a1)a^{-1} = aa^{-1} = 1.$$

Thus $cc^{-1} = cd$, so $c^{-1} = d$ by left cancellation. □

Definition. If $(G, *)$ is a finite group, with elements ordered $\{g_1, g_2, \ldots, g_n\}$, then the *table* for G (with respect to this ordering) is the $n \times n$ matrix whose (i, j) entry is $g_i * g_j$.

For example, the table for $\mathbb{Z}_4$ with respect to the standard ordering is

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

Fact. Suppose G is a *finite* group. In the table for its operation,

(1) Each row is a permutation of G.
(2) Each column is a permutation of G.

To see why, suppose that some element u occurs twice in row a, so

$\cdot$	1	$\cdots$	b	$\cdots$	c	$\cdots$
1	1	$\cdots$	b	$\cdots$	c	$\cdots$
$\vdots$						
a	a		u		u	
$\vdots$						

This means $ab = ac$, but $b \neq c$, contradicting left cancellation. Hence no element occurs more than once in the row labeled by a. Thus the function $f : G \to G$ given by $f : x \mapsto ax$ is injective. Since G is finite, it follows that f is a bijection, i.e., a permutation of G.

The same thing works for any row, and a similar argument works for any column.

5. Sept 16 – Isomorphisms, subgroups

Some jargon:

(1) G is **abelian** (named for Niels Henrik Abel, 1802–1829) if $ab = ba$ for all $a, b \in G$.
(2) If $a \in G$ then $\langle a \rangle$ denotes the set $\{a^n : n \in \mathbb{Z}\}$. Thus $\langle a \rangle \subseteq G$.
(3) G is **cyclic** if there exists $a \in G$ such that $G = \langle a \rangle$.
    - In this case we call a a **generator** of G.

Examples

(1) $(\mathbb{Z}, +)$, $(\mathbb{R}, +)$, $(\mathbb{Z}_n, + \pmod n)$, $(\mathbb{Z}_n^\times, \cdot \pmod n)$ are abelian.
(2) D_{2n}, S_n are not abelian ($n \geq 3$).
(3) Suppose $\mathbb{F}$ is a field (e.g., $\mathbb{F}$ could be $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$, or $\mathbb{Z}_p$ (p prime)). $GL_n(\mathbb{F})$ denotes the set of all invertible $n \times n$ matrices with entries from $\mathbb{F}$. $GL_n(\mathbb{F})$ with matrix multiplication is a group (exercise). It is nonabelian if $n \geq 2$.
(4) $(\mathbb{Z}, +)$ is cyclic, generated by 1. This is because $\langle 1 \rangle = \{n1 : n \in \mathbb{Z}\} = \mathbb{Z}$.
(5) Similarly, $(\mathbb{Z}_n, +)$ is cyclic for every $n \geq 1$.
(6) $(\mathbb{R}, +)$ is not cyclic. Are D_{2n}, S_n cyclic? No: every cyclic group is abelian (exercise).
(7) Is $(\mathbb{Z}_n^\times, \cdot)$ cyclic?

Section 1.6: Isomorphisms

The most fundamental relation between groups is that of **isomorphism**.

Definition. Let $\mathbb{G} = (G, \star)$ and $\mathbb{H} = (H, \diamond)$ be groups. A function $\varphi : G \to H$ is an *isomorphism from* $\mathbb{G}$ *to* $\mathbb{H}$ if it is a bijection and

$$\varphi(x \star y) = \varphi(x) \diamond \varphi(y) \qquad \text{for all } x, y \in G.$$

Example. Suppose $\mathbb{G} = (\mathbb{Z}_4, +)$ and $\mathbb{H} = (\mathbb{Z}_5^\times, \cdot)$.

(1) The map $\varphi : G \to H$ given by $\varphi(i) = i + 1$ is <u>not</u> an isomorphism. It is a bijection, but when $x = y = 3$, the requirement $\varphi(x + y) = \varphi(x) \cdot \varphi(y)$ fails because

$$\varphi(3 + 3) = \varphi(2) = 3 \qquad \text{while} \qquad \varphi(3) \cdot \varphi(3) = 4 \cdot 4 = 1.$$

(2) The map $\psi : G \to H$ given by $\psi(i) = 2^i \pmod 5$ <u>is</u> an isomorphism.
    (a)

i	0	1	2	3
$\psi(i)$	1	2	4	3

, so ψ is a bijection.
    (b) Regarding the condition $\psi(x + y) = \psi(x) \cdot \psi(y)$, note first that for all $a, b \in \mathbb{Z}$, if $a \equiv b \pmod 4$ then $2^a \equiv 2^b \pmod 5$. Now suppose $i, j, k \in \mathbb{Z}_4$ with $i + j = k$ (in $\mathbb{Z}_4$, meaning $i + j \equiv k \pmod 4$). Then $2^{i+j} \equiv 2^k \pmod 5$, so

$$\psi(i + j) = \psi(k) = 2^k \pmod 5 = 2^{i+j} \pmod 5 = 2^i \cdot 2^j \pmod 5 = \psi(i) \cdot \psi(j).$$

Exploring the last example, consider the tables for $(\mathbb{Z}_4, +)$ and for $(\mathbb{Z}_5^\times, \cdot)$. If we order $\mathbb{Z}_4$ as $(0, 1, 2, 3)$ and order $\mathbb{Z}_5^\times$ as $(1, 2, 4, 3)$, then the tables are

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

$\cdot$	1	2	4	3
1	1	2	4	3
2	2	4	3	1
4	4	3	1	2
3	3	1	2	4

The tables are "the same" modulo identifying $0 \mapsto 1$, $1 \mapsto 2$, $2 \mapsto 4$, and $3 \mapsto 3$; i.e., $x \mapsto \psi(x)$.

Theology:

(1) If φ is an isomorphism from $\mathbb{G}$ to $\mathbb{H}$, then the operation tables for $\mathbb{G}$ and $\mathbb{H}$ are "the same" (modulo the translation given by φ).
(2) If the operation tables for $\mathbb{G}$ and $\mathbb{H}$ are "the same" in this sense, then $\mathbb{G}$ and $\mathbb{H}$ are "essentially the same group."

Definition. We say that group $\mathbb{G}$ and $\mathbb{H}$ are *isomorphic* and write $\mathbb{G} \cong \mathbb{H}$ if there exists an isomorphism $\varphi : G \to H$.

Chapter 2

Definition. Let $\mathbb{G} = (G, \cdot)$ be a group. A *subgroup* of $\mathbb{G}$ is a subset $H \subseteq G$ satisfying

(1) $H \neq \varnothing$.
(2) H is closed under products; i.e., $a, b \in H$ implies $ab \in H$.
(3) H is closed under inverses; i.e., $a \in H$ implies $a^{-1} \in H$.

Example. Suppose $\mathbb{G} = (\mathbb{Z}, +)$.

(1) Let $E = \{\ldots, -4, -2, 0, 2, 4, \ldots\}$. Obviously nonempty. If $a, b \in E$ then a, b are even, so $a + b$ is even, so $a + b \in E$. Similarly, if a is even then so is $-a$, so E is closed under inverses. Thus H is a subgroup of $(\mathbb{Z}, +)$.
(2) The set of odd integers is *not* a subgroup of $(\mathbb{Z}, +)$.

6. Sept 18 – Cosets, Lagrange's Theorem

Fact: if $\mathbb{G} = (G, \cdot)$ is a group and H is a subgroup of $\mathbb{G}$, then $\mathbb{H} = (H, \cdot\restriction_H)$ is a group in its own right.

In particular, $1 \in H$. We can prove this as follows: pick any $a \in H$ (can do, $H \neq \varnothing$). Then $a^{-1} \in H$, so $a \cdot a^{-1} \in H$. ✓

Conventions.

(1) We also say that $\mathbb{H}$ is a subgroup of $\mathbb{G}$.
(2) We don't distinguish between $\mathbb{H}$ and H (or between $\mathbb{G}$ and G).
(3) We write $H \leq G$ to mean H is a subgroup of G.

Example. Let G be a group and $a \in G$. Recall that $\langle a \rangle = \{a^n : n \in \mathbb{Z}\}$.
Claim: $\langle a \rangle \leq G$.

Proof. Clearly $\langle a \rangle \subseteq G$ and $\langle a \rangle \neq \varnothing$. We check closure under the operation and under inverses.

(1) Given $a^n, a^m \in \langle a \rangle$, we have $a^n \cdot a^m = a^{n+m} \in \langle a \rangle$. ✓
(2) Given $a^n \in \langle a \rangle$, its inverse (in G) is $(a^n)^{-1} = a^{-n}$, which is in $\langle a \rangle$. ✓ □

$\langle a \rangle$ is called a **cyclic subgroup** of G **generated by** a.

Chapter 3

Definition. Suppose G is a group, $H \leq G$, and $a \in G$. The *left coset of H determined by a* is the set

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

E.g., $1H = H$. **Caution**: aH is generally *not* a subgroup of G.

Lemma. *For all $a \in G$, $|aH| = |H|$. Hence all left cosets of H have the same size as H.*

Proof. We define a bijection from H to aH in the only reasonable way. Define $f : H \to aH$ by $f(h) = ah$. f is surjective by definition. Suppose $f(h_1) = f(h_2)$. I.e., $ah_1 = ah_2$. Then $h_1 = h_2$ by left cancellation, so f is injective. Thus f is a bijection from H to aH. Hence $|H| = |aH|$. □

Caution: It can happen that $aH = bH$ even if $a \neq b$.

Proposition 3.4. *Suppose $H \leq G$. The set of left cosets of H partition G; that is,*

(1) $\bigcup\{aH : a \in G\} = G$
(2) *If $aH \neq bH$ then $aH \cap bH = \varnothing$.*

Proof. (1) Every $a \in G$ is an element of aH (since $1 \in H$), so clearly $G \subseteq \bigcup\{aH : a \in G\}$. The other inclusion is obvious.

(2) I'll show the contrapositive: $aH \cap bH \neq \varnothing$ implies $aH = bH$. Suppose $x \in aH \cap bH$. Thus $x = ah = bh'$ for some $h, h' \in H$. Then $a = xh^{-1} = (bh')h^{-1} = bh_1$ where $h_1 = h'h^{-1}$. Observe that $h \in H$ implies $h^{-1} \in H$, and with $h' \in H$ implies $h_1 = h'h^{-1} \in H$. As $a = bh_1$, this proves $a \in bH$.

Now let ah_2 be an arbitrary element of aH. Then $ah_2 = (bh_1)h_2 = bh_3 \in bH$. This proves $aH \subseteq bH$. A symmetric argument proves $bH \subseteq aH$. □

Theorem 3.8 (Lagrange's Theorem)**.** *Suppose G is a finite group and $H \leq G$. Then $|H|$ divides $|G|$.*

Proof. Let $|G| = n$ and $|H| = k$. H has only finitely many distinct left cosets; list them as $H, a_2H, \ldots, a_mH$. They partition G and all have the same size, namely k. Hence $n = mk$. □

Corollary. *If G is a finite group and $a \in G$, then $|\langle a\rangle|$ divides $|G|$.*

Proof. Because $\langle a\rangle \leq G$. □

On Assignment 2 you will show that $|\langle a\rangle| = \circ(a)$. Hence:

Corollary. *Suppose G is a finite group and $a \in G$. Then $\circ(a)$ divides $|G|$.*

Here are two nice applications.

Corollary 3.9. *If G is a finite group and $|G| = n$, then $x^n = 1$ for all $x \in G$.*

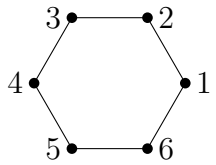
Proof. Given $x \in G$, let $k = \circ(x)$. Then $x^k = 1$. We have $k|n$ by the previous Corollary, say $n = km$. Then $x^n = x^{km} = (x^k)^m = 1^m = 1$. □

Corollary 3.10. *If G is a finite group and $|G| = p$ is prime, then G is cyclic.*

Proof. Pick any $a \in G$ satisfying $a \neq 1$. Then $\circ(a)|p$, so $\circ(a) = 1$ or p. If $\circ(a) = 1$ then $a^1 = 1$, contradicting $a \neq 1$. So $\circ(a) = p$. Hence $|\langle a\rangle| = p$. But $\langle a\rangle \subseteq G$ and $|G| = p$, which forces $\langle a\rangle = G$. So G is cyclic. □

7. Sept 22 – Cosets (continued), Normal subgroups

Consider the dihedral group D_{12} of symmetries of the regular hexagon C_6.



Also recall that $D_{12} = \{r^i : 0 \le i < 6\} \cup \{sr^i : 0 \le i \le 6\}$ where

- r = rotation ccw by $\pi/3$ radians = (1 2 3 4 5 6).
- s = reflection through the x-axis = (2 6)(3 5).
- $r^6 = s^2 = 1$ and $rs = sr^{-1}$.

What are some subgroups of D_{12}?

For starters, we know that $|D_{12}| = 12$, so by Lagrange's theorem, a subgroup can only have order 1, 2, 3, 4, 6 or 12.

Looking at cyclic subgroups, we find

$$\begin{aligned} \langle 1 \rangle &= \{1\} \\ \langle r \rangle &= \{1, r, r^2, r^3, r^4, r^5\} \\ \langle s \rangle &= \{1, s\} \\ \langle sr \rangle &= \{1, sr\}, \quad \text{etc. for any reflection} \end{aligned}$$

Can we find a subgroup of order 3? (Yes: $\langle r^2 \rangle$.)

Can we find a subgroup of order 4? (Yes: $\langle r^3, s \rangle$. This subgroup is not cyclic.)

Each subgroup has left cosets. For example:

- The left cosets of $\langle r \rangle$ are $\langle r \rangle = \{1, r, r^2, r^3, r^4, r^5\}$ and $s\langle r \rangle = \{s, sr, sr^2, sr^3, sr^4, sr^5\}$.
- Let $H = \langle s \rangle$. The left cosets of H are

$$\begin{aligned} H &= \{1, s\}, \\ rH &= \{r, rs\} = \{r, sr^5\}, \\ r^2H &= \{r^2, r^2s\} = \{r^2, sr^4\}, \\ r^3H &= \{r^3, r^3s\} = \{r^3, sr^3\}, \\ r^4H &= \{r^4, r^4s\} = \{r^4, sr^2\}, \\ r^5H &= \{r^5, r^5s\} = \{r^5, sr\}. \end{aligned}$$

Subgroups also have *right* cosets. For example, the right cosets of $H = \langle s \rangle$ are

$$\begin{aligned} H &= \{1, s\}, \\ Hr &= \{r, sr\}, \\ Hr^2 &= \{r^2, sr^2\}, \\ Hr^3 &= \{r^3, sr^3\}, \\ Hr^4 &= \{r^4, sr^4\}, \\ Hr^5 &= \{r^5, sr^5\}. \end{aligned}$$

Note that H has the same **number** of right and left cosets, but they aren't the same sets.

Definition. If G is a group and $H \leq G$, the **index of H in G**, denoted $[G : H]$, is the number of distinct left (or right) cosets of H.

Of course, if G is finite then $[G : H] = \frac{|G|}{|H|}$.

Definition. Suppose $H \leq G$. We say that H is a **normal subgroup**, and write $H \lhd G$, if $aH = Ha$ for all $a \in G$.

Of course if G is abelian then every subgroup is normal. In the example above, $\langle r \rangle \lhd D_{12}$ but $\langle s \rangle \not\lhd D_{12}$.

Definition. If $H, K \leq G$, then $HK := \{hk : h \in H \text{ and } k \in K\}$.

For example, $HH = \{h_1 h_2 : h_1, h_2 \in H\} = H$. In general, $H \subseteq HK$ and $K \subseteq HK$, but HK need not be a subgroup. For example, in D_{12} let $H = \langle s \rangle = \{1, s\}$ and $K = \langle sr \rangle = \{1, sr\}$. Then

$$\begin{aligned} HK &= \{1, s\}\{1, sr\} \\ &= \{1(1),\ 1(sr),\ s(1),\ s(sr)\} \\ &= \{1,\ sr,\ s,\ r\}. \end{aligned}$$

This isn't a subgroup because $r, s \in HK$ but $rs \notin HK$.

Proposition. *Suppose G is a group and $H, K \leq G$. If either $H \lhd G$ or $K \lhd G$, then $HK \leq G$.*

Proof. Assume $H \lhd G$. I'll first show $HK = KH$. Indeed,

$$HK = \bigcup_{k \in K} Hk = \bigcup_{k \in K} kH = KH.$$

Now I'll show $HK \leq G$. Certainly $HK \subseteq G$ and $HK \neq \varnothing$.

- Given $a, b \in HK$, write $a = h_1k_1$ and $b = h_2k_2$ with $h_1, h_2 \in H$ and $k_1, k_2 \in K$. We have $k_1h_2 \in KH$, so $k_1h_2 \in HK$, so $k_1h_2 = h_3k_3$ for some $h_3 \in H$ and $k_3 \in K$. Then
$$(h_1k_1)(h_2k_2) = h_1(k_1h_2)k_2 = h_1(h_3k_3)k_2 = (h_1h_3)(k_3k_2) \in HK$$
since $H, K \leq G$. This proves closure under $\cdot$.
- Suppose $a \in HK$, say $a = hk$. Then $a^{-1} = k^{-1}h^{-1} \in KH = HK$, proving closure under inverses. □

Definition. Suppose G is a group and $H \leq G$. The *normalizer* of H, denoted $N_G(H)$, is the set
$$N_G(H) = \{a \in G : aH = Ha\}.$$

Note that $H \lhd G$ iff $N_G(H) = G$. Also note that the proof of the previous Proposition didn't use the full assumption that $H \lhd G$; it only needed $kH = Hk$ for all $k \in K$, or equivalently, $K \subseteq N_G(H)$. Hence:

Corollary 3.15. *Suppose G is a group and $H, K \leq G$. If $K \subseteq N_G(H)$ (or $H \subseteq N_G(K)$), then $HK \leq G$.*

8. Sept 23 – Direct products

Recall that we write $H \lhd G$ if $aH = Ha$ for all $a \in G$. The following is a useful characterization.

Theorem 3.6. *Suppose $H \leq G$. TFAE:*

(1) $H \lhd G$.
(2) $aHa^{-1} = H$ *for all* $a \in G$.
(3) $aHa^{-1} \subseteq H$ *for all* $a \in G$.
(4) *If* $h \in H$, *then* $aha^{-1} \in H$ *for all* $a \in G$.

Proof. (1) $\Leftrightarrow$ (2) $\Rightarrow$ (3) $\Leftrightarrow$ (4) is obvious. To finish, it suffices to prove (3) $\Rightarrow$ (2). Assume (3). In particular, $a^{-1}Ha = a^{-1}H(a^{-1})^{-1} \subseteq H$. Multiply on left by a and on right by a^{-1} to get $H \subseteq aHa^{-1}$. With (3) this gives $H = aHa^{-1}$ for all $a \in G$, giving (2). □

Here is a cute application.

Lemma. *Suppose $H, K \lhd G$ and $H \cap K = \{1\}$. Then $hk = kh$ for all $h \in H$ and $k \in K$.*

Proof. We use the following trick. Given $a, b \in G$, their *commutator* is $[a, b] = a^{-1}b^{-1}ab$. It is easy to show that $ab = ba$ iff $[a, b] = 1$ (exercise). So now let $h \in H$ and $k \in K$ and consider $[h, k] = h^{-1}k^{-1}hk$. We can write

$$[h, k] = h^{-1}(k^{-1}hk).$$

Since $h \in H \lhd G$ and $k^{-1} \in G$ we get $k^{-1}hk \in H$, say $k^{-1}hk = h_1$. Then

$$[h, k] = h^{-1}h_1 \in H.$$

Similarly,

$$[h, k] = (\underbrace{h^{-1}k^{-1}h}_{\in K})k \in K.$$

Hence $[h, k] \in H \cap K = \{1\}$, proving $[h, k] = 1$, so $hk = kh$. □

Let $(G_1, \star)$ and $(G_2, \diamond)$ be groups. Their *direct product* is $(G_1 \times G_2, *)$ where

$$(a_1, a_2) * (b_1, b_2) \;\; = \;\; (a_1 \star b_1, a_2 \diamond b_2).$$

Fact: If G_1, G_2 are groups, then $G_1 \times G_2$ is also a group.

Proof sketch. Let's check existence of inverses. (The identity element is $\mathbf{1} = (1_1, 1_2)$ where 1_i is the identity element of G_i.) For any $\mathbf{a} = (a_1, a_2) \in G_1 \times G_2$, I claim that $\mathbf{a}^{-1} := (a_1^{-1}, a_2^{-1})$ is an inverse to $\mathbf{a}$.

$$\mathbf{a} * \mathbf{a}^{-1} = (a_1, a_2) * (a_1^{-1}, a_2^{-1}) \stackrel{\text{df}}{=} (a_1 \star a_1^{-1},\, a_2 \diamond a_2^{-1}) = (1_1, 1_2) = \mathbf{1}.$$

A similar proof shows $\mathbf{a}^{-1} * \mathbf{a} = \mathbf{1}$, so we're good. □

Notation.

(1) If both $\star$ and $\diamond$ are written as $+$, then we may write $*$ as $+$.
(2) Products of more factors are defined analogously. $G^n = \underbrace{G \times G \times \cdots \times G}_{n}$.

Consider $(\mathbb{Z}_2, +)^2$. It has 4 elements: $(0,0)$, $(0,1)$ $(1,0)$, $(1,1)$. Its table:

$+$	(0,0)	(1,0)	(0,1)	(1,1)
(0,0)	(0,0)	(1,0)	(0,1)	(1,1)
(1,0)	(1,0)	(0,0)	(1,1)	(0,1)
(0,1)	(0,1)	(1,1)	(0,0)	(1,0)
(1,1)	(1,1)	(0,1)	(1,0)	(0,0)

If we rename its elements $1, b, c, d$ we get the table

$\cdot$	1	b	c	d
1	1	b	c	d
b	b	1	d	c
c	c	d	1	b
d	d	c	b	1

We recognize this from Assignment 1; it is the table for $\mathbb{Z}_8^\times$. Hence $\mathbb{Z}_8^\times \cong (\mathbb{Z}_2, +)^2$. We have **factored** $\mathbb{Z}_8^\times$.

Theorem 5.9. *Let G be a group. Suppose there exist $H, K \lhd G$ satisfying*

(1) $H \cap K = \{1\}$;
(2) $HK = G$.

Then $G \cong H \times K$.

Proof. Define a function $\varphi : H \times K \to G$ by $\varphi((h,k)) = hk$. Let's prove that φ is an isomorphism. First check that it is a bijection. It is surjective because $HK = G$. We'll prove that it's injective using $H \cap K = \{1\}$. Indeed, suppose $\varphi((h_1, k_1)) = \varphi((h_2, k_2))$, i.e., $h_1 k_1 = h_2 k_2$. Multiply on left by h_2^{-1} and on right by k_1^{-1} to get $h_2^{-1} h_1 = k_2 k_1^{-1}$. The left side is in H while the right side is in K, so both sides are in $H \cap K = \{1\}$, proving $h_2^{-1} h_1 = 1 = k_2 k_1^{-1}$. These equations imply $h_1 = h_2$ and $k_1 = k_2$, proving φ is injective.

[To be continued]

9. Sept 25 – Homomorphisms

Proof of Theorem (continued). Recall that G is a group, $H, K \lhd K$, $H \cap K = \{1\}$ and $HK = G$. The claim is that $G \cong H \times K$. I've defined $\varphi : H \times K \to G$ by $\varphi((h,k)) = hk$ and shown that it is a bijection.

Finally we must prove that

$$\varphi((h_1,k_1) * (h_2,k_2)) = \varphi((h_1,k_1)) \cdot \varphi((h_2,k_2)),$$

i.e.,

$$\varphi((h_1h_2, k_1k_2)) = \varphi((h_1,k_1)) \cdot \varphi((h_2,k_2)),$$

i.e., $(h_1h_2)(k_1k_2) = (h_1k_1)(h_2k_2)$. This follows because $h_2k_1 = k_1h_2$ by yesterday's cute lemma. □

Example. Let $G = (\mathbb{Z}_6, +)$. Let $H = \langle 2 \rangle = \{0,2,4\}$ and $K = \langle 3 \rangle = \{0,3\}$. Clearly

- $H, K \leq \mathbb{Z}_6$. (Cyclic subgroups are subgroups)
- $H, K \lhd \mathbb{Z}_6$. (Because $\mathbb{Z}_6$ is abelian)
- $H \cap K = \{0\}$.
- $H + K = \{h + k : h \in \{0,2,4\} \text{ and } k \in \{0,3\}\} = \mathbb{Z}_6$.

Hence by the Theorem, $(\mathbb{Z}_6, +) \cong H \times K$.

It is also easy to see that $H \cong (\mathbb{Z}_3, +)$. (The isomorphism $\mathbb{Z}_3 \to H$ is the map $x \mapsto 2x$.) Similarly, $K \cong (\mathbb{Z}_2, +)$. Hence $(\mathbb{Z}_6, +) \cong (\mathbb{Z}_3, +) \times (\mathbb{Z}_2, +)$.

A similar argument shows $(\mathbb{Z}_{mn}, +) \cong (\mathbb{Z}_m, +) \times (\mathbb{Z}_n, +)$ provided $\gcd(m,n) = 1$.

Definition. Let $G = (G, \star)$ and $H = (H, \diamond)$ be groups. A function $\varphi : G \to H$ is a *homomorphism* if

$$\varphi(x \star y) = \varphi(x) \diamond \varphi(y) \qquad \text{for all } x, y \in G.$$

Example.

(1) Any isomorphism is a homomorphism.

(2) The parity function $\text{par} : \mathbb{Z} \to \{0,1\}$ given by

$$\text{par}(n) = \begin{cases} 0 & \text{if } n \text{ is even} \\ 1 & \text{if } n \text{ is odd} \end{cases}$$

is a homomorphism from $(\mathbb{Z}, +)$ to $(\mathbb{Z}_2, +)$, because $\text{par}(x+y)$, i.e., the parity of $x + y$, is the mod-2 sum of $\text{par}(x)$ and $\text{par}(y)$.

(3) More generally, the function $\mathbb{Z} \to \mathbb{Z}_n$ given by $x \mapsto (x \pmod n)$ is a homomorphism.

(4) Let $G = (\mathbb{C}^\times, \cdot)$ and $H = (\mathbb{R}^\times, \cdot)$. The function $\varphi : \mathbb{C}^\times \to \mathbb{R}^\times$ given by $\varphi(z) = |z|$ is a homomorphism, because $\varphi(zw) = |zw| = |z||w| = \varphi(z)\varphi(w)$.

Definition 9.1. Let $\varphi : G \to H$ be a homomorphism.

(1) $\mathrm{im}(\varphi)$ denotes the *image* (or *range*) of φ.
(2) The *kernel* of φ is the set

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

It is denoted $\ker \varphi$.

Proposition 3.1. *Let $\varphi : G \to H$ be a homomorphism.*

(1) $\varphi(1_G) = 1_H$.
(2) $\varphi(g^{-1}) = \varphi(g)^{-1}$ *for all* $g \in G$.
(3) *More generally,* $\varphi(g^n) = \varphi(g)^n$ *for all* $n \in \mathbb{Z}$.
(4) $\ker \varphi \leq G$ *and* $\mathrm{im}(\varphi) \leq H$.

Proof. (1) Pick $g \in G$. We have $\varphi(1_G)\varphi(g) = \varphi(1_G \cdot g) = \varphi(g) = 1_H \cdot \varphi(g)$. Right cancellation in H gives $\varphi(1_G) = 1_H$.

(2) Let $h = \varphi(g)$. Then $h\varphi(g^{-1}) = \varphi(g)\varphi(g^{-1}) = \varphi(gg^{-1}) = \varphi(1_G) = 1_H = hh^{-1}$. Left cancellation gives $\varphi(g^{-1}) = h^{-1} = \varphi(g)^{-1}$.

(3) Exercise.

(4) Clearly $\ker \varphi \subseteq G$ and $\mathrm{im}(\varphi) \subseteq H$. Since $\varphi(1_G) = 1_H$ by (1) we have $1_G \in \ker \varphi$ and $1_H \in \mathrm{im}(\varphi)$, so $\ker \varphi, \mathrm{im}(\varphi) \neq \varnothing$. Remains to check closure under products and inverses.

First $\ker \varphi$:

- Suppose $a, b \in \ker \varphi$, meaning $\varphi(a) = \varphi(b) = 1_H$. Then $\varphi(ab) = \varphi(a)\varphi(b) = 1_H \cdot 1_H = 1_H$, proving $ab \in \ker \varphi$.
- Suppose $a \in \ker \varphi$, meaning $\varphi(a) = 1_H$. Then $\varphi(a^{-1}) = \varphi(a)^{-1}$ (by (2)) $= (1_H)^{-1} = 1_H$, proving $a^{-1} \in \ker \varphi$.

Next $\mathrm{im}(\varphi)$:

- Suppose $x, y \in \mathrm{im}(\varphi)$. So there exist $a, b \in G$ with $x = \varphi(a)$ and $y = \varphi(b)$. Then $xy = \varphi(a)\varphi(b) = \varphi(ab)$. As $ab \in G$, this proves $xy \in \mathrm{im}(\varphi)$.
- Suppose $x \in \mathrm{im}(\varphi)$, say $x = \varphi(a)$ where $a \in G$. Then $x^{-1} = \varphi(a)^{-1} = \varphi(a^{-1})$ by (2). Since $a^{-1} \in G$, this proves $x^{-1} \in \mathrm{im}(\varphi)$.

□

More is true:

Proposition. *Let $\varphi : G \to H$ be a homomorphism. Then* $\ker \varphi \lhd G$.

Proof. It suffices to show $g \in \ker \varphi$ implies $aga^{-1} \in \ker \varphi$ for all $a \in G$. So suppose $g \in \ker \varphi$ and $a \in G$. To show $aga^{-1} \in \ker \varphi$, we evaluate it under φ:

$$\begin{aligned} \varphi(aga^{-1}) &= \varphi(a)\varphi(ga^{-1}) = \varphi(a)\varphi(g)\varphi(a^{-1}) = \varphi(a) \cdot 1_H \cdot \varphi(a^{-1}) \quad (\text{because } g \in \ker \varphi) \\ &= \varphi(a) \cdot \varphi(a^{-1}) = \varphi(a \cdot a^{-1}) = \varphi(1_G) = 1_H. \end{aligned}$$

Hence $aga^{-1} \in \ker \varphi$. □

10. Sept 29 – Quotient groups

(Section 3.1, but not following the text closely.)

Definition. Suppose G is a group and $H \leq G$. Then G/H denotes the set of all *left* cosets of N.

Examples.

(1) If $G = (\mathbb{Z}, +)$ and $H = \langle 5 \rangle = 5\mathbb{Z}$, then $G/H = \mathbb{Z}/5\mathbb{Z} = \{5\mathbb{Z}, 1 + 5\mathbb{Z}, 2 + 5\mathbb{Z}, 3 + 5\mathbb{Z}, 4 + 5\mathbb{Z}\}$.
(2) If $G = (\mathbb{C}^\times, \cdot)$ and $H = S = \{z : |z| = 1\}$, then $G/H = \mathbb{C}^\times/S = \{\text{all circles centred at } 0\}$.

In general, we want to define an operation $\cdot$ on G/H. That is, given two left cosets C, D of H, we want to define another left coset $C \cdot D$. The most natural choice is

$$C \cdot D \stackrel{\text{df}}{=} CD = \{cd : c \in C,\, d \in D\},$$

or equivalently, $(aH) \cdot (bH) = (aH)(bH)$. There is a problem: $(aH)(bH)$ might not be a left coset of H. However there is no problem when H is normal.

Proposition 3.5 (1)**.** *Suppose $N \lhd G$.*

(1) *$aN, bN \in G/N$ implies $(aN)(bN) \in G/N$.*
(2) *$(aN)(bN) = (ab)N$ for all $a, b \in G$.*

Proof. $(aN)(bN) = a(Nb)N = a(bN)N$ (by normality) $= (ab)NN = (ab)N$, which is a left coset of N. □

Definition. If $N \lhd G$, then $\cdot$ is defined on G/N by $(aN) \cdot (bN) \stackrel{\text{df}}{=} (ab)N$.

Notation. If the operation of G is $+$, then we write $+$ instead of $\cdot$ for the operation on G/N as well. In this case the definition is $(a + N) + (b + N) \stackrel{\text{df}}{=} (a + b) + N$.

Example. In $\mathbb{C}^\times/S$, let $C = \{z : |z| = r\}$ and $D = \{z : |z| = s\}$. What is $C \cdot D$?

Solution. Then $C = rS$ and $D = sS$, so

$$\begin{aligned} C \cdot D &= (rS)(sS) \\ &= \{(re^{i\theta})(se^{i\varphi}) : \theta, \varphi \in \mathbb{R}\} \\ &= \{rse^{i(\theta+\varphi)} : \theta, \varphi \in \mathbb{R}\} \\ &= \{rse^{i(\psi)} : \psi \in \mathbb{R}\} \\ &= (rs)S. \end{aligned}$$

Example. Let $G = (\mathbb{Z}_5, +)$ and $N = 5\mathbb{Z}$. If $C = 3 + 5\mathbb{Z}$ and $D = 4 + 5\mathbb{Z}$, what is $C + D$?

Solution. $(3 + 5\mathbb{Z}) + (4 + 5\mathbb{Z}) = (3 + 4) + 5\mathbb{Z} = 7 + 5\mathbb{Z} = 2 + 5\mathbb{Z}$.

Proposition 3.5 (2)**.** *Suppose $N \lhd G$. Then $(G/N, \cdot)$ is a group.*

Proof. We've already seen that $\cdot$ is an operation on G/N. What remains is to prove that $\cdot$ is associative, has a 2-sided identity element, and every left coset of N has a 2-sided inverse (with respect to $\cdot$).

Associativity: For all $aN, bN, cN \in G/N$,

$$\begin{aligned} aN \cdot (bN \cdot cN) &= aN \cdot (bc)N \\ &= a(bc)N \\ &= (ab)cN \\ &= (ab)N \cdot cN \\ &= (aN \cdot bN) \cdot cN. \end{aligned}$$

Identity: Obviously $N = 1N \in G/N$. We will show that N is an identity element with respect to $\cdot$. For any $aN \in G/N$,

$$aN \cdot N \;=\; aN \cdot 1N \;=\; (a1)N \;=\; aN.$$

Similarly, $N \cdot aN = aN$.

Inverses: For any $aN \in G/N$, of course we have $a^{-1}N \in G/N$. We will show that $a^{-1}N$ is an inverse to aN.

$$\begin{aligned} aN \cdot a^{-1}N &= (aa^{-1})N \\ &= 1N \\ &= N \end{aligned}$$

and similarly $a^{-1}N \cdot aN = N$.

□

Definition. Suppose $N \lhd G$. The group $(G/N, \cdot)$ is called the **quotient group of** G **by** N (or of G **modulo** N).

Examples.

(1) $\mathbb{Z}/5\mathbb{Z}$. Though complicated (its elements are the 5 cosets of $5\mathbb{Z}$), this quotient group is easily seen to be isomorphic to $\mathbb{Z}_5$. The isomorphism $\mathbb{Z}_5 \to \mathbb{Z}/5\mathbb{Z}$ sends $a \mapsto a + 5\mathbb{Z}$. (In fact, most textbooks define $\mathbb{Z}_5$ to *be* $\mathbb{Z}/5\mathbb{Z}$.)
(2) More generally, $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$.
(3) $\mathbb{C}^\times/S$. One can show that this quotient group is isomorphic to $(\mathbb{R}^{>0}, \cdot)$ in the obvious way.
(4) Let N be the subgroup of D_{12} given by $N = \langle r^3 \rangle = \{1, r^3\}$. One can check (by tedious calculations) that $N \lhd D_{12}$. Hence we can form the quotient group D_{12}/N. Its elements are the left cosets of N in D_{12}. What is D_{12}/N isomorphic to?

11. Sept 30 – 1st Isomorphism theorem

Section 3.3

Definition. Suppose $N \lhd G$. Define $\pi_N : G \to G/N$ by $\pi_N(g) = gN$.

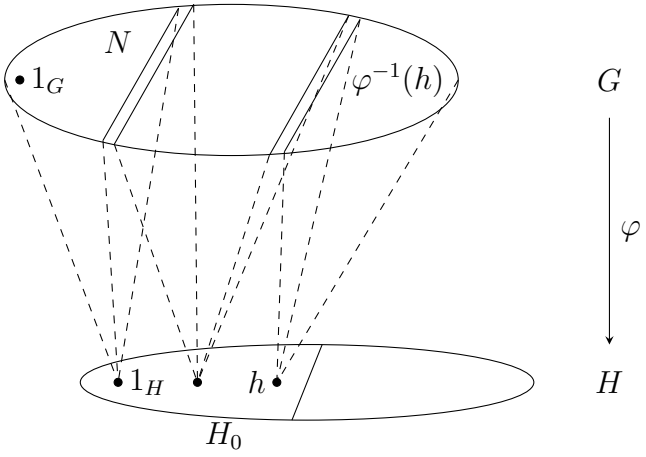
π_N is called the "mod N projection map."

Lemma. *If $N \lhd G$, then $\pi_N : G \to G/N$ is a homomorphism and* $\ker \pi_N = N$.

Proof. $\pi_N(ab) = (ab)N = (aN)(bN) = \pi_N(a)\pi_N(b)$, proving π_N is a homomorphism. $\ker \pi_N = \{g \in G : \pi_N(g) = 1_{G/N}\} = \{g \in G : gN = N\} = N$. □

Consider now an arbitrary homomorphism $\varphi : G \to H$. In general we can't assume that φ is injective or surjective. Let $N = \ker \varphi$ and $H_0 = \text{im}(\varphi)$ and recall that $N \lhd G$ and $H_0 \leq H$.

For each $h \in H_0$, the preimage $\varphi^{-1}(h) := \{g \in G : \varphi(g) = h\}$ is called the *fiber* of φ above h. Note that the fiber above 1_H is $\ker \varphi = N$.



Proposition 3.2. *Suppose $\varphi : G \to H$ is a homomorphism and $N = \ker \varphi$. The fibers of φ are precisely the (left) cosets of N.*

Proof. Let $h \in \text{im}(\varphi)$ and choose $a \in \varphi^{-1}(h)$. I will show that $\varphi^{-1}(h) = aN$.

$\boxed{\varphi^{-1}(h) \subseteq aN.}$ Let $b \in \varphi^{-1}(h)$. Then

$$\varphi(a^{-1}b) = \varphi(a^{-1})\varphi(b) = \varphi(a)^{-1}\varphi(b) = h^{-1}h = 1_H.$$

Hence $a^{-1}b \in N$. So $b = a(a^{-1}b) \in aN$.

$\boxed{aN \subseteq \varphi^{-1}(h).}$ Let $x \in aN$, so $x = an$ for some $n \in N$. Then

$$\varphi(x) = \varphi(an) = \varphi(a)\varphi(n) = h \cdot 1_H = h$$

so $x \in \varphi^{-1}(h)$. □

Observe that φ is injective iff each of its fibers consists of just one element. By the previous result, this holds iff $|N| = 1$. This proves:

Corollary 3.17. *A homomorphism $\varphi : G \to H$ is injective iff* $\ker\varphi = \{1_G\}$.

We are ready for our second important theorem (the first was Lagrange's theorem).

Theorem 3.16 (1st Isomorphism Theorem)**.** *Suppose $\varphi : G \to H$ is a surjective homomorphism. Then $G/\ker\varphi \cong H$.*

Proof. Let $N = \ker\varphi$. Define $\overline{\varphi} : G/N \to H$ by the rule $\overline{\varphi}(aN) = \varphi(a)$. $\overline{\varphi}$ will be our isomorphism. We must first check that this is well-defined: i.e., if $aN = bN$ do we have $\varphi(a) = \varphi(b)$? Yes: if $aN = bN$, then a, b belong to the same coset of N, so they belong to the same fiber of φ by Proposition 3.2, meaning $\varphi(a) = \varphi(b)$.

Next we check that $\overline{\varphi}$ is a homomorphism. The question is whether, for any $aN, bN \in G/N$, we have

$$\overline{\varphi}(aN \cdot bN) \stackrel{?}{=} \overline{\varphi}(aN)\overline{\varphi}(bN).$$

Well, $\overline{\varphi}(aN \cdot bN) = \overline{\varphi}((ab)N) = \varphi(ab) = \varphi(a)\varphi(b) = \overline{\varphi}(aN)\overline{\varphi}(bN)$. So $\overline{\varphi}$ is a homomorphism.

Clearly $\overline{\varphi}$ is surjective (because φ is). It remains only to check that $\overline{\varphi}$ is injective, or equivalently, that $\ker\overline{\varphi} = \{N\}$. Suppose $aN \in G/N$. Then

$$aN \in \ker\overline{\varphi} \iff \overline{\varphi}(aN) = 1_H \iff \phi(a) = 1_H \iff a \in N \iff aN = N.$$

Hence $\ker\overline{\varphi} = \{N\} = \{1_{G/N}\}$, proving $\overline{\varphi}$ is injective. In summary, $\overline{\varphi} : G/N \cong H$. □

Examples.

(1) Define $\varphi : \mathbb{C}^\times \to \mathbb{R}^{>0}$ by $\varphi(z) = |z|$. We've already seen that this is a homomorphism, and it is easy to see that it is surjective. Thus by the 1st Isomorphism Theorem, $\mathbb{C}^\times/\ker\varphi \cong \mathbb{R}^{>0}$. What is $\ker\varphi$? Clearly $\ker\varphi = \{z \in \mathbb{C}^\times : |z| = 1\} = S$, the unit circle. Thus $\mathbb{C}^\times/S \cong \mathbb{R}^{>0}$.

(2) Recall that $D_{12} = \{r^i : 0 \le i < 6\} \cup \{sr^i : 0 \le i < 6\}$ where r is a counter-clockwise rotation by 60° and s is the reflection through the x-axis. Let's write $D_6 = \{t^i : 0 \le i < 3\} \cup \{st^i : 0 \le i < 3\}$ where t is the counter-clockwise rotation by 120°. Then clearly $t = r^2$, so $D_6 = \{1, r^2, r^4\} \cup \{s, sr^2, sr^4\}$.

Now define a function $\varphi : D_{12} \to D_6$ by $\varphi(r^i) = r^{2i}$ and $\varphi(sr^i) = sr^{2i}$. One can check (by a tedious consideration of cases) that $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in D_{12}$; hence φ is a homomorphism. Clearly φ is surjective, and it's easy to calculate that $\ker\varphi = \{1, r^3\}$.

It follows by the 1st Isomorphism Theorem that $D_{12}/\{1, r^3\} \cong D_6$.

12. Oct 2 – 2nd and 3rd Isomorphism theorems

In the proof of the 1st Isomorphism theorem, recall that $\overline{\varphi}$ satisfied (in fact was defined by)

$$\overline{\varphi}(aN) = \varphi(a) \qquad \text{for all } a \in G.$$

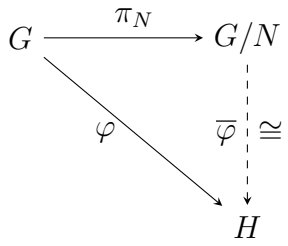
This can be restated as

$$\overline{\varphi}(\pi_N(a)) = \varphi(a) \qquad \text{for all } a \in G$$

which is equivalent to

$$\overline{\varphi} \circ \pi_N \;=\; \varphi.$$

We say that φ **factors through** π_N via φ. Pictorially,



Applications

(1) Given any group, we can define $\varphi : G \to G$ by $\varphi(x) = x$. It is easy to see that φ is a surjective homomorphism and $\ker\varphi = \{1\}$. Hence by the 1st Isomorphism Theorem,

$$G/\{1\} \cong G.$$

(2) Suppose we have a group G, a normal subgroup $N \lhd G$, and another subgroup $H \leq G$. We can form G/N. Define a function $\varphi : H \to G/N$ by $\varphi(a) = aN$. It is easy to check that φ is a homomorphism: $a, b \in H$ implies $\varphi(ab) = (ab)N = (aN)(bN) = \varphi(a)\varphi(b)$. But φ need not be onto.

What is $\mathrm{im}(\varphi)$? As a ranges over H, aN ranges over the left cosets of N in HN. Hence $\mathrm{im}(\varphi) = HN/N$.

Incidentally, we know that $HN \leq G$ because $N \lhd G$ (Sept 22), so HN is a group. Clearly $N \leq HN$. Now

$$\begin{aligned} N \lhd G \quad &\iff \quad gNg^{-1} = N \text{ for all } g \in G \\ &\implies \quad gNg^{-1} = N \text{ for all } g \in HN \\ &\iff \quad N \lhd HN. \end{aligned}$$

Thus $N \lhd HN$, so HN/N is in fact a group. Thus φ is a **surjective** homomorphism from H to HN/N.

By the 1st Isomorphism Theorem, $HN/N \cong H/\ker\varphi$. So we calculate $\ker\varphi$:

$$\ker\varphi = \{h \in H \,:\, hN = N\} = \{h \in H \,:\, h \in N\} = H \cap N.$$

Hence $HN/N \cong H/(H \cap N)$. This proves:

Theorem 3.18 (2nd Isomorphism Theorem)**.** *Suppose G is a group and $H, N \leq G$ with $N \lhd G$. Then $HN/N \cong H/(H \cap N)$.*

(3) Now suppose G is a group and we have two normal subgroups $N, K \lhd G$ with $N \leq K$. We can form G/N and G/K. Define $\varphi : G/N \to G/K$ by $\varphi(aN) = aK$. We have to check that this is well-defined: if $aN = bN$, then $N = a^{-1}bN$, so $a^{-1} \in N$. This implies $a^{-1}b \in K$, so $a^{-1}bK = K$, so $bK = aK$. This proves φ is well-defined.

Obviously φ is surjective.

We check that φ is a homomorphism: for all $aN, bN \in G/N$,

$$\varphi((aN)\cdot(bN)) = \varphi((ab)N) = (ab)K = (aK)(bK) = \varphi(aN)\varphi(bN).$$

Thus φ is a surjective homomorphism from G/N to G/K. Hence by the 1st Isomorphism Theorem, $G/K \cong (G/N)/\ker\varphi$. What is $\ker\varphi$? Calculate:

$$\ker\varphi = \{aN \in G/N \,:\, aK = K\} = \{aN \,:\, a \in K\} = K/N.$$

This proves:

Theorem 3.19 (3rd Isomorphism Theorem)**.** *Suppose G is a group and $N, K \lhd G$ with $N \leq K$. Then $K/N \lhd G/N$ and $(G/N)/(K/N) \cong G/K$.*

(4) Finally, suppose G is a group and $N, K \lhd G$ satisfy $NK = G$ and $N \cap K = \{1\}$. Recall that we know that this implies $G \cong N \times K$. What can we say about G/N and G/K?

By the 2nd Isomorphism theorem, $NK/K \cong N/(N \cap K)$. Since $NK = G$ and $N \cap K = \{1\}$, this gives $G/N \cong K/\{1\} \cong K$. By symmetry, $G/K \cong N$.

13. Oct 6 – Group actions

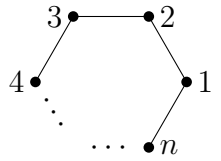
Sections 1.7 and 4.1

Definition. Let G be a group and X a set. An **action of G on X** is a map $g \mapsto \pi_g$ which assigns to each $g \in G$ a permutation $\pi_g \in S_X$, and which "respects the operation of G," in the sense that if $g, h \in G$ then $\pi_{gh} = \pi_g \circ \pi_h$.

In other words, an action of G on X is a homomorphism $\pi : G \to S_X$.

Examples.

(1) S_X acts naturally on X via the map $\sigma \mapsto \sigma$.
(2) D_{2n} acts naturally on the set $\{1, 2, \ldots, n\}$ via the picture



For example, in D_{12} we have $\pi_r = (1\ 2\ 3\ 4\ 5\ 6)$ and $\pi_s = (2\ 6)(3\ 5)$, and

$$\pi_{sr} = (1\ 6)(2\ 5)(3\ 4) = (2\ 6)(3\ 5)(1\ 2\ 3\ 4\ 5\ 6) = \pi_s \circ \pi_r.$$

In general, we naturally have a map $\pi : D_{2n} \to S_n$ and it is a homomorphism.

(3) G naturally acts on itself in a number of ways. Here is one of them. For each $g \in G$ let $\psi_g : G \to G$ given by $\psi_g(x) = gxg^{-1}$. ψ_g is an automorphism of G so is certainly a permutation of G. Moreover, for any $g, h \in G$,

$$\begin{aligned} \psi_{gh}(x) &= (gh)x(gh)^{-1} = (gh)xh^{-1}g^{-1} = g(hxh^{-1})g^{-1} \\ &= g\psi_h(x)g^{-1} = \psi_g(\psi_h(x)) = (\psi_g \circ \psi_h)(x), \end{aligned}$$

true for all $x \in G$, so $\psi_{gh} = \psi_g \circ \psi_h$. Hence the map $g \mapsto \psi_g$ is a homomorphism $\psi : G \to S_G$.

Notation. If π is an action of G on X, and $g \in G$ and $a, b \in X$, then we express $\pi_g(a) = b$ by writing $g \cdot a = b$ and say *g moves a to b*. Note that $\pi_{gh} = \pi_g \circ \pi_h$ translates to $(gh) \cdot a = g \cdot (h \cdot a)$ for all $a \in X$. Note also that $\pi_1 = \mathrm{id}$, which translates to $1 \cdot a = a$ for all $a \in X$.

Definition. Let π be an action of G on X.

(1) The **kernel** of the action is the kernel of π as a homomorphism $G \to S_X$.
(2) The action is **faithful** if its kernel is $\{1\}$ (equivalently, if π is injective).
(3) Given $a \in X$, the **orbit** of a is the set $G \cdot a = \{g \cdot a \ : \ g \in G\}$ of places to which a gets moved by elements of G.

Note: if G acts faithfully on X, then G is isomorphic to a subgroup of S_X. (π is the isomorphism.)

Proposition 4.2 (1)**.** *Suppose G acts on X. The orbits of the action partition X.*

Proof. As $a \in G \cdot a$ for each $a \in X$, the orbits clearly cover X. Suppose $G \cdot a, G \cdot b$ are two orbits with $G \cdot a \cap G \cdot b \neq \varnothing$. Pick $x \in G \cdot a \cap G \cdot b$; thus pick $g, h \in G$ with $g \cdot a = x = h \cdot b$. Then

$$(h^{-1}g) \cdot a = h^{-1} \cdot (g \cdot a) = h^{-1} \cdot (h \cdot b) = (h^{-1}h) \cdot b = 1 \cdot b = b,$$

proving $b \in G \cdot a$. It is easy to show that $b \in G \cdot a$ implies $G \cdot b \subseteq G \cdot a$. Dually we get $G \cdot a \subseteq G \cdot b$. So $G \cdot a = G \cdot b$. □

Definition. An action of G on X is **transitive** if it has only one orbit (X).

Example. The natural action of D_{2n} on $\{1, 2, \ldots, n\}$ is transitive, since for any $i, j \in \{1, \ldots, n\}$ we can find $g \in D_{2n}$ such that $g \cdot i = j$.

Definition. Let π be an action of G on X. Given $a \in X$, the **stabilizer** of a is the set

$$G_a = \{g \in G \,:\, g \cdot a = a\}.$$

Proposition 4.2 (2)**.** *Suppose G acts on X. For every $a \in X$:*

(1) $G_a \leq G$.
(2) $|G \cdot a| = [G : G_a]$.

Hence if G is finite, then every orbit has size dividing $|G|$.

Proof. (1) Exercise.

(2) Recall that $G \cdot a = \{g \cdot a \,:\, g \in G\}$. Observe that for $g, h \in G$,

$$g \cdot a = h \cdot a \iff h^{-1} \cdot (g \cdot a) = h^{-1} \cdot (h \cdot a)$$
$$\iff (h^{-1}g) \cdot a = a \iff h^{-1}g \in G_a \iff hG_a = gG_a.$$

In other words, the $g \cdot a$ depends only on gG_a. Thus the number of distinct values of $g \cdot a$ equals the number of distinct left cosets of G_a in G. □

Example. Consider the natural action of $G = D_{2n}$ on $\{1, 2, \ldots, n\}$. This action is transitive. Thus for any $i \in \{1, 2, \ldots, n\}$, $|G \cdot i| = n$. Hence by the Orbit-Stabilizer Theorem, the stabilizer of i must have order 2.

14. Oct 7 – Permutation representations and Cayley's Theorem

Let G be a group. G acts on itself by left multiplication: $g \cdot a = ga$. Call this action λ; so for $g \in G$, λ_g is the permutation in S_G given by $\lambda_g(a) = ga$.

Note that λ is a homomorphism $G \to S_G$, since for any $g, h \in G$,

$$\lambda_{gh}(a) = (gh)a = g(ha) = \lambda_g(\lambda_h(a)) = (\lambda_g \circ \lambda_h)(a) \qquad \text{for all } a \in G,$$

proving $\lambda_{gh} = \lambda_g \circ \lambda_h$.

Note also that if $g \neq 1$ then $\lambda_g(1) = g1 = g \neq 1$, so $\lambda_g \neq \mathrm{id}$. Hence $\ker \lambda = \{1\}$, so the action is faithful. Hence G is isomorphic to a subgroup of S_G. This proves

Cayley's Theorem (Corollary 4.4) Every group is isomorphic to a subgroup of a symmetric group. If $|G| = n$, then G is isomorphic to a subgroup of S_n.

We proved Cayley's theorem by exhibiting a faithful action of G on a $|G|$-element set (namely, G itself). Sometimes we can find smaller sets on which G faithfully acts.

Example. Suppose $\{1\} < H < G$. Recall that G/H is the set of left cosets of H. G acts on G/H by left multiplication:

$$g \cdot aH = (ga)H.$$

It is easy to check that this is an action (i.e., the map $\lambda^H : G \to S_{G/H}$ where λ^H_g is the permutation $aH \mapsto (ga)H$ is a homomorphism). Let $N = \ker \lambda^H$. Note that if $g \in N$ then $\lambda^H_g = \mathrm{id}$, so $\lambda^H_g(H) = H$, meaning $gH = H$, which implies $g \in H$. Thus $N \subseteq H$.

Proposition. *Suppose G is a finite group, $\{1\} < H < G$, and G has no normal subgroups contained in H except for $\{1\}$. Then G is isomorphic to a subgroup of S_m where $m = [G : H]$.*

Proof. Let λ^H be the action of G on G/H by left multiplication. The kernel of λ^H is a normal subgroup of G contained in H, so must be $\{1\}$. Hence λ^H is faithful, so $\lambda : G \cong \mathrm{Im}(\lambda^H) \leq S_{G/H} \cong S_m$. □

The action of G on G/H is also the key to the proof of the following.

Corollary 4.5. *Suppose G is a finite group and p is the smallest prime dividing $|G|$. If $H \leq G$ with $[G : H] = p$, then $H \lhd G$.*

Proof. Let λ^H be the action of G on G/H by left multiplication. Let $N = \ker \lambda^H$. Thus $N \lhd G$ and $N \subseteq H$. Whether or not λ^H is injective, we know from the 1st Isomorphism Theorem that $G/N \cong \mathrm{im}(\lambda^H) \leq S_{G/H} \cong S_p$, so G/N is isomorphic to a subgroup of S_p. Hence $|G/N|$ divides $|S_p| = p!$ by Lagrange's theorem. What else can we say about $|G/N|$?

Because $N \subseteq H$, we can define $[H : N] = k$ and get

$$|G/N| = \frac{|G|}{|N|} = \frac{|G|}{|H|} \cdot \frac{|H|}{|N|} = pk.$$

Thus $pk|p!$, and hence $k|(p-1)!$ But k divides $|H|$, which divides $|G|$, so every prime divisor of k must be $\geq p$ (by choice of p). This forces $k = 1$, which implies $N = H$, so $H \lhd G$ (as $N \lhd G$). □

Here is one final result of this kind. Suppose G is a group with $|G| = 52$, and $H \leq G$ with $|H| = 13$. I claim that these assumptions imply $H \lhd G$. Here's why. Let λ^H be the action of G on G/H. Let $N = \ker \lambda^H$. Then $N \lhd G$ and $N \subseteq H$. The latter fact implies $|N|$ divides $|H| = 13$ (by Lagrange), so $|N| = 1$ or 13.

Suppose $|N| = 1$. Then λ^H is faithful, so $\lambda^H : G \cong \mathrm{Im}(\lambda^H) \leq S_{[G/H]}$, proving G is isomorphic to a subgroup of S_4. But $|G| = 52$ while $|S_4| = 4! = 24$, so G can't possibly be isomorphic to a subgroup of S_4. This case is impossible.

Hence $|N| = 13$. But $N \subseteq H$ and $|H| = 13$. These facts imply $N = H$. As $N \lhd G$, it follows that $H \lhd G$.

15. Oct 9 – Class equation and Cauchy's theorem

Let's explore the action of G on itself by conjugation: ψ_g is the map $x \mapsto gxg^{-1}$. Thus $g \cdot a = gag^{-1}$. Given $a \in G$, what is the orbit $G \cdot a$?

$$G \cdot a = \{gag^{-1} \,:\, g \in G\}, \quad \text{the set of conjugates of } a.$$

Let's denote this set by $\mathrm{Conj}(a)$. It is called the **conjugacy class** of a.

Example. If $G = S_3$, then

$$\begin{aligned} \mathrm{Conj}(\mathrm{id}) &= \{\pi \circ \mathrm{id} \circ \pi^{-1} \,:\, \pi \in S_3\} = \{\mathrm{id}\} \\ \mathrm{Conj}((1\ 2)) &= \{\pi \circ (1\ 2) \circ \pi^{-1} \,:\, \pi \in S_3\} = \{(1\ 2), (1\ 3), (2\ 3)\} \\ \mathrm{Conj}((1\ 2\ 3)) &= \{\pi \circ (1\ 2\ 3) \circ \pi^{-1} \,:\, \pi \in S_3\} = \{(1\ 2\ 3), (1\ 3\ 2)\}. \end{aligned}$$

By general properties of actions, we get:

(1) G is partitioned by its conjugacy classes.
(2) For any $a \in G$, $|\mathrm{Conj}(a)| = [G : G_a]$.

In particular, $|\mathrm{Conj}(a)|$ divides $|G|$ when G is finite.

What is the stabilizer G_a?

$$G_a = \{g \in G \,:\, g \cdot a = a\} = \{g \in G \,:\, gag^{-1} = a\} = \{g \in G \,:\, ga = ag\} = C_G(a).$$

Thus

Proposition 4.6. In any group G, $|\mathrm{Conj}(a)| = [G : C_G(a)]$.

Of particular interest are the 1-element conjugacy classes. Observe that

$$\begin{aligned} \mathrm{Conj}(a) = \{a\} &\iff |\mathrm{Conj}(a)| = 1 \\ &\iff [G : C_G(a)] = 1 \\ &\iff C_G(a) = G \\ &\iff ga = ag \text{ for all } g \in G \\ &\iff a \in Z(G). \end{aligned}$$

Hence

Proposition. *Every group G is the disjoint union of $Z(G)$ and its nontrivial conjugacy classes.*

This fact is called the *class equation*. It has useful consequences.

Theorem 4.7. If p is a prime and $|G| = p^n$, then $Z(G) \neq \{1\}$.

Proof. By the class equation,

$$p^n = |G| = |Z(G)| + \sum_{i=1}^{m} |\mathrm{Conj}(a_i)|$$

where $\mathrm{Conj}(a_1), \ldots, \mathrm{Conj}(a_m)$ are the nontrivial conjugacy classes. Each $|\mathrm{Conj}(a_i)|$ divides p^n and is > 1. Hence $|Z(G)| \equiv 0 \pmod{p}$, so $|Z(G)| \geq p$. □

Cauchy's Theorem. *Suppose G is a finite group. If p is prime and p divides $|G|$, then there exists an element in G of order p.*

Proof. Note that it suffices to prove the existence of $a \in G$ such that $p|\circ(a)$. (If $\circ(a) = pk$, then $\circ(a^k) = p$.)

We first prove this weaker claim for *abelian* groups, by induction on $|G|$.

Base: $|G| = p$. Then G is cyclic of order p, done.

Inductive step: $|G| = pm$, $m > 1$. Pick any element $a \in G \setminus \{1\}$.

Case 1: p divides $\circ(a)$. Then we're done.

Case 2: p does not divide $\circ(a)$.

Let $N = \langle a \rangle$. Clearly $\{1\} < N < G$. Also, $N \lhd G$ (as G is abelian) so we get G/N. $|N| > 1$ implies $|G/N| < |G|$. Also $|G| = |N| \cdot |G/N|$ so p divides $|G/N|$. In fact, G/N is an *abelian* group (exercise), so the inductive hypothesis applies and we get an element $bN \in G/N$ of order p.

Let $n = \circ(b)$. Then $b^n = 1$, so $(bN)^n = N$, so $p|n$, so we're done.

Now we prove the general case, again by induction on G. Look at the class equation:

$$|G| = |Z(G)| + \sum_{i=1}^{m} |\mathrm{Conj}(a_i)|.$$

Case 1. For some i we have that $|\mathrm{Conj}(a_i)|$ is *not* divisible by p.

Since $|\mathrm{Conj}(a_i)| = \frac{|G|}{|C_G(a_i)|}$ it must be that p divides $C_G(a_i)$. Note that $C_G(a_i) \leq G$, and $C_G(a_i) \neq G$ (as $a_i \notin Z(G)$). So we can apply the inductive hypothesis to $C_G(a_i)$ to get an element of order p.

Case 2. p divides every $|\mathrm{Conj}(a_i)|$.

Then p divides $|Z(G)|$. Note that $Z(G)$ is abelian so our earlier argument gives an element of order p. □

16. Oct 14 – Finite abelian groups

Lemma. *Suppose G is an abelian group and $m \in \mathbb{Z}$. Define $G^{(m)} = \{a \in G : a^m = 1\}$. Then $G^{(m)} \leq G$.*

Proof. $a, b \in G^{(m)} \implies a^m = b^m = 1 \implies (ab)^m = a^m b^m = 1 \implies ab \in G^{(m)}$. Similarly, $a \in G^{(m)} \implies a^{-1} \in G^{(m)}$. □

Lemma. *Suppose G is finite abelian, $|G| = mk$ with $\gcd(m,k) = 1$. Then*

1. *$G \cong G^{(m)} \times G^{(k)}$.*
2. *$|G^{(m)}| = m$ and $|G^{(k)}| = k$.*

Proof. (1) First, suppose $a \in G^{(m)} \cap G^{(k)}$. Then $a^m = a^k = 1$. Pick $x, y \in \mathbb{Z}$ with $mx + ky = 1$. Then

$$a = a^{mx+ky} = (a^m)^x \cdot (a^k)^y = 1^x \cdot 1^y = 1.$$

This proves $\boxed{G^{(m)} \cap G^{(k)} = \{1\}}$.

Next, note that if $a \in G$ then $1 = a^{mk} = (a^k)^m = (a^m)^k$, proving $a^k \in G^{(m)}$ and $a^m \in G^{(k)}$. Hence

$$a = a^{mx+ky} = \underbrace{(a^k)^y}_{\in G^{(m)}} \cdot \underbrace{(a^m)^x}_{\in G^{(k)}} \in G^{(m)} \cdot G^{(k)}.$$

This proves $\boxed{G = G^{(m)} \cdot G^{(k)}}$. Obviously $\boxed{G^{(m)}, G^{(k)} \lhd G}$. So $G \cong G^{(m)} \times G^{(k)}$.

(2) Let $|G^{(m)}| = m'$ and $|G^{(k)}| = k'$. Clearly $mk = |G| = m'k'$ by (1). Suppose $\gcd(m, k') \neq 1$. Then we can choose a prime p dividing both m and k'. By Cauchy's theorem, there exists $a \in G^{(k)}$ with $\circ(a) = p$. But $p|m$ implies $a^m = 1$, so $a \in G^{(m)}$. Hence $a = 1$, contradiction. This proves $\gcd(m, k') = 1$.

Now $m|m'k'$ and $\gcd(m, k') = 1$ imply $m|m'$ and hence $m \leq m'$. A similar argument gives $k \leq k'$. As $mk = m'k'$, this can only happen if $m = m'$ and $k = k'$. □

Corollary. *Suppose G is finite abelian and $|G| = n = p_1^{n_1} \cdots p_k^{n_k}$ where $p_1, \ldots, p_k$ are distinct primes.*

1. *$G \cong G^{(p_1^{n_1})} \times \cdots \times G^{(p_k^{n_k})}$*
2. *$|G^{(p_i^{n_i})}| = p_i^{n_i}$ for each i.*

This is called the **primary decomposition** of G.

Example. Let $G = \mathbb{Z}_{13}^\times$. $|G| = 12 = 2^2 \cdot 3$.

$$\begin{aligned} G^{(4)} &= \{a \in \mathbb{Z}_{13}^\times : a^4 = 1\} = \{1, 5, 8, 12\} \\ G^{(3)} &= \{a \in \mathbb{Z}_{13}^\times : a^3 = 1\} = \{1, 3, 9\} \end{aligned}$$

By the Lemma, $\{1,5,8,12\}\{1,3,9\} = \mathbb{Z}_{13}^\times \cong \{1,5,8,12\} \times \{1,3,9\}$.

Definition. A finite group is a p-**group** if $|G| = p^n$ for some $n \geq 1$.

17. Oct 16 – Finite abelian groups (continued)

On Tuesday we saw that every finite abelian group can be factored as a direct product of finite abelian p-groups (i.e., of order p^n where p is prime). Today we will see how to factor finite abelian p-groups.

Definition. Let G be an abelian group. A **basis** for G is a sequence $a_1, \ldots, a_t \in G$ satisfying

(1) $a_i \neq 1$ for all i.
(2) $G = \langle a_1 \rangle \langle a_2 \rangle \cdots \langle a_t \rangle$.
(3) For all $i < t$,
$$\langle a_1 \rangle \langle a_2 \rangle \cdots \langle a_i \rangle \cap \langle a_{i+1} \rangle = \{1\}.$$

Suppose $a_1, \ldots, a_t$ is a basis for G. Let $N_i = \langle a_i \rangle$ and $H_i = N_1 N_2 \cdots N_i$. Condition (3) says $H_i \cap N_{i+1} = \{1\}$ for all $i < t$. Note that $H_i N_{i+1} = H_{i+1}$. Since G is abelian we of course have $H_i, N_{i+1} \lhd H_{i+1}$. Hence $H_{i+1} \cong H_i \times N_{i+1}$ (for each $i < t$). Thus

$$\begin{aligned} G &= H_t && \text{by (2)} \\ &\cong H_{t-1} \times N_t \\ &\cong (H_{t-2} \times N_{t-1}) \times N_t \\ &\vdots \\ &\cong N_1 \times N_2 \times \cdots \times N_t. \end{aligned}$$

Thus

Proposition. *If G is abelian and G has a basis, then G is isomorphic to a direct product of cyclic groups.*

Theorem. *Every finite abelian p-group has a basis.*

Proof sketch. Let $|G| = p^n$. Start be choosing a_1 to be an element of G with $\circ(a_1)$ maximum, say $\circ(a_1) = p^{n_1}$. If $n_1 = n$ then $G = \langle a_1 \rangle$ and we're done. Otherwise, let $H_1 = \langle a_1 \rangle$, form G/H_1, and pick an element $bH_1 \in G/H_1$ with $\circ(bH_1)$ maximum. Observe that $|G/H| = p^n/p^{n_1} = p^{n-n_1}$, so $\circ(bH_1) = p^{n_2}$ for some $n_2 \leq n - n_1$.

Also let $\circ(b) = p^k$. Then $k \leq n_1$ (by choice of a_1). Furthermore, $b^{p^k} = 1$, so $(bH_1)^{p^k} = H$, so $p^{n_2} | p^k$, so $n_2 \leq k$. Thus $n_2 \leq k \leq n_1$.

Claim. *There exists $a_2 \in bH_1$ with $\circ(a_2) = p^{n_2}$.*

Proof. From $(bH)^{p^{n_2}} = H$ we get $b^{p^{n_2}} \in H_1 = \langle a_1 \rangle$. Write $b^{p^{n_2}} = a_1^i$. Then

$$a_1^{ip^{n_1-n_2}} = (b^{p^{n_2}})^{p^{n_1-n_2}} = b^{p^{n_1}}.$$

Note that $k \leq n_1$ (proved above), so $\circ(b) = p^k | p^{n_1}$, so $\boxed{b^{p^{n_1}} = 1}$. Hence $a_1^{ip^{n_1-n_2}} = 1$, which implies $\circ(a_1) | ip^{n_1-n_2}$, i.e., $p^{n_1} | ip^{n_1-n_2}$, which implies $p^{n_2} | i$, say $i = jp^{n_2}$. Now define $a_2 = ba_1^{-j} \in bH_1$. Note that:

(1) $(a_2)^{p^{n_2}} = (ba_1^{-j})^{p^{n_2}} = b^{p^{n_2}}(a_1^{jp^{n_2}})^{-1} = b^{p^{n_2}}(a_1^i)^{-1} = 1$.
(2) If $0 \leq t < p^{n_2}$ and $a_2^t = 1$, so $(ba_1^{-j})^t = 1$, then $b^t \in \langle a_1 \rangle = H_1$, so $(bH_1)^t = H_1$. As $\circ(bH_1) = p^{n_2}$, this forces $t = 0$.

Together, these facts prove $\circ(a_2) = p^{n_2}$ as claimed. □

Claim. $H_1 \cap \langle a_2 \rangle = \{1\}$.

Proof. Assume $0 \leq t < p^{n_2}$ and $a_2^t \in H_1$. The same calculation proving item (2) above shows $b^t \in H_1$, so $(bH_1)^t = H_1$, so $t = 0$. Thus $H_1 \cap \{1, a_2, a_2^2, \ldots, a_2^{p^{n_2}-1}\} = \{1\}$. □

Now we let $H_2 = \langle a_1 \rangle \langle a_2 \rangle$, If $H_2 = G$ then we're done. Otherwise, form G/H_2, and pick an element $cH_2 \in G/H_2$ with $\circ(cH_2)$ maximum. We get $\circ(cH_2) = p^{n_3}$ for some $n_3 \leq n-(n_1+n_2)$. We can prove $c^{p^{n_1}} = 1$ and $(cH_1)^{p^{n_2}} = H_1$ by our choice of a_1, bH_1. (For example, $\circ(cH_1) = p^\ell$ for some ℓ, and $\circ(cH_1) \leq \circ(bH_1) = p^{n_2}$ by our choice of bH_1, so $\ell \leq n_2$, so $p^\ell | p^{n_2}$, so $(cH_1)^{p^{n_2}} = H_1$.) Hence $\boxed{c^{p^{n_1}} = 1}$ and $\boxed{c^{p^{n_2}} \in H_1}$.

Claim. *There exists $a_3 \in cH_2$ with $\circ(a_3) = p^{n_3}$.*

Proof sketch. $(cH_2)^{p^{n_3}} = H_2 = \langle a_1 \rangle \langle a_2 \rangle$, so $c^{p^{n_3}} = a_1^{i_1} a_2^{i_2}$ for some i_1, i_2. The two boxed facts above can be shown to imply $p^{n_3} | i_1$ and $p^{n_3} | i_2$ respectively.

(Details: recall that $c^{p^{n_2}} \in H_1$. But

$$c^{p^{n_2}} = (c^{p^{n_3}})^{p^{n_2-n_3}} = (a_1^{i_1} a_2^{i_2})^{p^{n_2-n_3}} = (a_1^{i_1 p^{n_2-n_3}})(a_2^{i_2 p^{n_2-n_3}}).$$

Hence $a_2^{i_2 p^{n_2-n_3}} = c^{p^{n_2}}(a_1^{i_1 p^{n_2-n_3}})^{-1} \in H_1$ as both factors are in H_1. This implies $(a_2 H_1)^{i_2 p^{n_2-n_3}} = H_1$, so $\circ(a_2 H_1) = p^{n_2} | i_2 p^{n_2-n_3}$, which in turn implies $p^{n_3} | i_2$. Next we prove $p^{n_3} | i_1$. Start with the fact, proved above, that $c^{p^{n_1}} = 1$. This implies

$$1 = c^{p^{n_1}} = (c^{p^{n_3}})^{p^{n_1-n_3}} = (a_1^{i_1} a_2^{i_2})^{p^{n_1-n_3}} = (a_1^{i_1 p^{n_1-n_3}})(a_2^{i_1 p^{n_1-n_3}}),$$

which implies

$$a_1^{i_1 p^{n_1-n_3}} = (a_2^{i_1 p^{n_1-n_3}})^{-1} \in \langle a_2 \rangle.$$

Obviously $a_1^{i_1 p^{n_1-n_3}} \in \langle a_1 \rangle = H_1$. So $a_1^{i_1 p^{n_1-n_3}} \in H_1 \cap \langle a_2 \rangle$, implying $a_1^{i_1 p^{n_1-n_3}} = 1$. But $\circ(a_1) = p^{n_1}$, so $p^{n_1} | i_1 p^{n_1-n_3}$, implying $p^{n_3} | i_1$.)

Now let $i_1 = j_1 p^{n_3}$ and $i_2 = j_2 p^{n_3}$ and define $a_3 = c a_1^{-j_1} a_2^{-j_2}$. This works. □

Claim. $H_2 \cap \langle a_3 \rangle = \{1\}$.

Proved similarly to $H_1 \cap \langle a_2 \rangle = \{1\}$.

Now if $G = H_2 \langle a_3 \rangle$ then we can stop. Otherwise, we must look for a_4. Start by letting $H_3 = H_2 \langle a_3 \rangle$ and forming G/H_3. Choose $(dH_3) \in G/H_3$ with $\circ(dH_3)$ maximum ...

And so on. As G is finite, this process must eventually stop, at which point $a_1, a_2, \ldots, a_t$ will be a basis for G. □

18. Oct 20 – Definition of a ring

Definition. A **ring** is an ordered triple $(R, +, \cdot)$ where

- R is a non-empty set;
- $+$ and $\cdot$ are binary operations on R;

which jointly satisfy the following conditions:

(i) $(R, +)$ is an abelian group;
(ii) $\cdot$ is associative;
(iii) There exists $1 \in R$ such that $1 \cdot a = a \cdot 1 = a$ for all $a \in R$.
(iv) (Distributive laws): for all $a, b, c \in R$,

$$\begin{aligned} (a+b)\cdot c &= (a\cdot c)+(b\cdot c) \\ a\cdot(b+c) &= (a\cdot b)+(a\cdot c). \end{aligned}$$

Notation/jargon.

- We denote $(R, +, \cdot)$ by R.
- The identity element of $(R, +)$ is denoted 0.
- The inverse of a in the group $(R, +)$ is denoted $-a$, and is called the *additive* inverse.
- We write $a - b$ for $a + (-b)$.
- The element 1 is called the *multiplicative identity*. It is (provably) unique.
- We say that R is *commutative* if it satisfies $a \cdot b = b \cdot a$ for all $a, b \in R$, and is *noncommutative* otherwise.

Example.

(1) $\mathbb{Z}$ (with usual addition and multiplication) is a commutative ring; it is the prototypical example of a commutative ring.
(2) $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are also commutative rings.
(3) For every $n \geq 2$, the set $M_n(\mathbb{R})$ of all $n \times n$ matrices over $\mathbb{R}$ (with matrix addition and multiplication) is a noncommutative ring. Similarly, $M_n(\mathbb{Z})$, $M_n(\mathbb{Q})$, $M_n(\mathbb{C})$.
(4) $\mathbb{Z}_n$ is a (finite) commutative ring for every $n \geq 2$.
(5) Let $C(\mathbb{R})$ be the set of all continuous, everywhere-defined functions $f : \mathbb{R} \to \mathbb{R}$ (a very big set). Define $f + g$ and $f \cdot g$ pointwise; that is,

$$\begin{aligned} (f+g)(x) &:= f(x)+g(x) \\ (f\cdot g)(x) &:= f(x)\cdot g(x). \end{aligned}$$

Then $(C(\mathbb{R}), +, \cdot)$ is a commutative ring. What is its zero element? Its identity element?
(6) Is $(C(\mathbb{R}), +, \circ)$ a ring?

Warning: In general, you cannot assume that $\cdot$ satisfies left or right cancellation. For example in $\mathbb{Z}$ we have $0 \cdot x = 0 \cdot y$, but that does not imply $x = y$. In some rings, even if $a \neq 0$, one cannot assume that $a \cdot b = a \cdot c$ implies $b = c$.

Proposition 7.1. *Let R be a ring. Then*

(1) $0a = a0 = 0$ *for all* $a \in R$.
(2) $-a = (-1)a = a(-1)$ *for all* $a \in A$.
(3) $(-a)b = a(-b) = -(ab)$ *for all* $a, b \in R$.
(4) $(-a)(-b) = ab$.

Proof. (1) $0 + 0 = 0$, so $0a = (0 + 0)a = (0a) + (0a)$. Hence $0a + 0 = 0a + 0a$, so cancelling (in the group $(R, +)$) gives $0a = 0$. Similar proof works for $a0 = 0$.

(2) $1 + (-1) = 0$. Hence $0 = 0a = (1 + (-1))a = (1a) + (-1)a = a + (-1)a$. Hence $a + (-a) = a + (-1)a$, so cancelling gives $(-1)a = -a$. Similar proof works for $a(-1) = -a$.

(3) $(-a)b = (a(-1))b = a((-1)b) = a(-b)$. Also, $(-a)b = ((-1)a)b = (-1)(ab) = -(ab)$.

(4) Exercise. □

Definition. Let R be a ring.

(1) An element $a \in R$ is a *unit* if there exists $b \in R$ satisfying $ab = ba = 1$. (We also say that a is *invertible*. b is called the *inverse* of a and is denoted a^{-1}; it is provably unique.)
(2) $R^\times$ denotes the set of units in R.

Remark. 2 is a unit in $\mathbb{Q}$ but is not a unit in $\mathbb{Z}$. $\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}$ while $\mathbb{Z}^\times = \{1, -1\}$.

In general, $(R^\times, \cdot)$ is a group; called the *group of units* of R.

Can $0 = 1$ in a ring? If $0 = 1$, then $a = a1 = a0 = 0$ for all $a \in R$, i.e., $R = \{0\}$. A 1-element ring is called *trivial*. Thus a ring is nontrivial iff it satisfies $0 \neq 1$.

Definition.

(1) A *division ring* is a ring D satisfying $0 \neq 1$ and $D^\times = D \setminus \{0\}$ (i.e., every nonzero element is a unit).
(2) A *field* is a commutative division ring.

Example.

(1) $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_p$ (p prime) are fields.

19. Oct 21 – Integral domains, subrings

Notation. Let R be a ring and $a \in R$.

(1) For $n > 1$ we let na denote $\underbrace{a + a + \cdots + a}_{n}$.
(2) For $n \geq 1$ we let $(-n)a$ denote $-(na)$. (Thus na is defined for all $n \in \mathbb{Z}$.)
(3) $\mathbb{Z}a = \{na : n \in \mathbb{Z}\}$.

Note that $\mathbb{Z}a$ is the cyclic subgroup of $(R, +)$ generated by a.

Recall:

Definition.

(1) A *division ring* is a ring D satisfying $0 \neq 1$ and $D^\times = D \setminus \{0\}$ (i.e., every nonzero element is a unit).
(2) A *field* is a commutative division ring.

Is $M_2(\mathbb{R})$ a division ring?

Example. $\mathbb{H}$, the *ring of real Hamiltonion quaternions*, is the set of all expressions $a + bi + dj + dk$ where $a, b, c, d \in \mathbb{R}$ and i, j, k are primitive symbols.

(1) Addition is defined obviously:

$$(a + bi + cj + dk) + (a' + b'i + c'j + d'k) = (a + a') + (b + b')i + (c + c')j + (d + d')k$$

(2) Multiplication is first defined on the primitive symbols:

$$i^2 = j^2 = k^2 = -1, \qquad ij = k,\ jk = i,\ ki = j, \qquad ji = -k,\ kj = -i,\ ik = -j.$$

(3) Then multiplication is extended to expressions by assuming $ai = ia$, $aj = ja$, $ak = ka$ for all $a \in \mathbb{R}$, and assuming distributivity.

It can be shown that $\mathbb{H}$ is a ring. If $a + bi + cj + dk \neq 0$ then

$$(a + bi + cj + dk)^{-1} = \frac{a}{e} - \frac{b}{e}i - \frac{c}{e}j - \frac{d}{e}k$$

where $e = a^2 + b^2 + c^2 + d^2$, so $\mathbb{H}$ is a division ring. $\mathbb{H}$ is not a field (as $ij \neq ji$).

Definition. Let R be a ring. A *zero divisor* is an element $a \in R$ such that

(1) $a \neq 0$, and
(2) There exists $b \in R$ with $b \neq 0$ such that either $ab = 0$ or $ba = 0$.

Example.

(1) $\mathbb{Z}$ has no zero divisors (since $ab = 0$ implies $a = 0$ or $b = 0$).
(2) $2, 3, 4$ are zero divisors in $\mathbb{Z}_6$, since $2 \cdot 3 = 4 \cdot 3 = 0$.
(3) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is a zero divisor in $M_2(\mathbb{R})$, since $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Proposition 7.2. *Suppose R is a ring and $a \in R$ with $a \neq 0$. If a is not a zero divisor, then we can "cancel by a." That is, for all $b, c \in R$,*

$$ab = ac \implies b = c$$
$$ba = ca \implies b = c.$$

Proof. Assume $ab = ac$. Then $a(b-c) = a(b+(-c)) = ab + a(-c) = (ab) + -(ac) = ab - ac = ab - ab = 0$. Since $a \neq -$ but a is not a zero divisor, it must be that $b - c = 0$, i.e., $b = c$. The other implication is proved similarly. □

Lemma. *If R is a ring and $a \in R^\times$, then a is not a zero divisor. Hence we can always "cancel by units."*

Proof. Argue by contradiction. Assume $a \in R^\times$ and a is a zero divisor. Thus a^{-1} exists in R, and there exists $b \in R$ with $b \neq 0$ such that either $ab = 0$ or $ba = 0$. Suppose $ba = 0$; then $b = b1 = b(aa^{-1}) = (ba)a^{-1} = 0a^{-1} = 0$, contradiction. The equation $ab = 0$ also leads to a contradiction. □

Definition. A ring R is called an *integral domain* (or *domain*) if it is commutative, satisfies $0 \neq 1$, and has no zero divisors.

For example, $\mathbb{Z}$ is an integral domain.

Corollary 7.3. *Every field is an integral domain.*

Proof. Follows from the previous lemma. □

Definition. Suppose R is a ring. A *subring* of R is a subset $S \subseteq R$ such that

(1) S is a subgroup of $(R, +)$.
(2) S is closed under multiplication (i.e., $a, b \in S$ implies $ab \in S$).
(3) $1 \in S$.

Write $S \leq R$ to denote that S is a subring of R.

As was the case for groups, every subring of a ring is itself a ring (with operations inherited from the larger ring).

Example.

(1) $\mathbb{Z} \leq \mathbb{Q}$. $\mathbb{R} \leq \mathbb{C}$.
(2) $\mathbb{Z}_n \not\leq \mathbb{Z}$.
(3) Is $M_2(\mathbb{R})$ a subring of $M_3(\mathbb{R})$? (No.)
(4) Is $M_2(\mathbb{Z})$ a subring of $M_2(\mathbb{R})$? (Yes.)
(5) Recall $C(\mathbb{R})$, the ring of all continuous functions $\mathbb{R} \to \mathbb{R}$. Define $P(\mathbb{R})$ to be the set of all *polynomial* functions (i.e., defined by polynomials, say with real coefficients). $P(\mathbb{R}) \leq C(\mathbb{R})$.

20. Oct 27 – Polynomial rings

Section 7.2

Let R be a ring. Let x be a formal variable.

- A **polynomial in x over R** is an expression

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

  where $n \geq 0$, $a_0, \ldots, a_n \in R$, and if $n > 0$ then $a_n \neq 0$.
- We also denote this by $\sum_{i=0}^{n} a_i x^i$.
- Note that if $n = 0$ then the expression is just a_0. When $n = 0$ and $a_0 = 0$ the expression is just 0. (This is the **zero polynomial**.)
- An expression $a_i x^i$ is called a **term** of the polynomial.
- The elements $a_0, a_1, \ldots, a_n$ of R are called the **coefficients** of the polynomial.
- The **degree** of the polynomial is n, except for the zero polynomial which has no degree.
- If the polynomial is not 0, then the **leading term** is $a_n x^n$, and the **leading coefficient** is a_n.
- By definition, two polynomials are equal iff they have the same degree and the same coefficients.

Definition. $R[x]$ denotes the set of all polynomials in x over R.

Key fact: every $p(x) \in R[x]$ can be viewed as a **formula** which defines a **function** $p : R \to R$. However, the polynomial is **not** the same thing as the function it defines.

- For example, if $R = \mathbb{Z}_2$, then $\mathbb{Z}_2[x]$ is an infinite set, but there are only 4 different polynomial functions $\mathbb{Z}_2 \to \mathbb{Z}_2$.

For convenience, we also define **sloppy polynomials over R** to be **all** expressions of the form $\sum_{i=0}^{n} a_i x^i$ $(a_0, \ldots, a_n \in R)$. Here the degree is n even if $a_n = 0$. Every sloppy polynomial determines a unique "tidy" polynomial (by deleting zero terms).

Definition. Given a ring R, define $+$ and $\cdot$ on $R[x]$ "in the obvious way." That is, given $p(x), q(x) \in R[x]$:

(1) To define $p(x) + q(x)$:
Write $p(x)$ and $q(x)$ as sloppy polynomials of the same degree and use

$$\left(\sum_{i=0}^{n} a_i x^i\right) + \left(\sum_{i=0}^{n} b_i x^i\right) \;=\; \sum_{i=0}^{n} (a_i + b_i) x^i.$$

(2) To define $p(x) \cdot q(x)$:

Write $p(x) = \sum_{i=1}^{m} a_i x^i$ and $q(x) = \sum_{j=0}^{n} b_j x^j$. Then

$$\begin{aligned}\left(\sum_{i=0}^{m} a_i x^i\right) \cdot \left(\sum_{j=0}^{n} b_j x^j\right) &= (a_0 + a_1 x + a_2 x^2 + \cdots) \cdot (b_0 + b_1 x + b_2 x^2 + \cdots) \\ &= (a_0 b_0) + (a_0 b_1 + a_1 b_0)x + (a_0 b_2 + a_1 b_1 + a_2 b_0)x^2 + \cdots \\ &= \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i b_j\right) x^k.\end{aligned}$$

(In both formulas, the right-hand sides must be trimmed if sloppy.)

Theorem. *$R[x]$ is a ring containing R as a subring.*

Proof sketch. A nightmare. To illustrate, I will prove that $\cdot$ is associative. Let $p(x) = \sum_i a_i x^i$, $q(x) = \sum_j b_j x^j$, and $r(x) = \sum_k c_k x^k$. Then $p(x) \cdot q(x) = \sum_s d_s x^s$ where

$$d_s = \sum_{i+j=s} a_i b_j,$$

so $(p(x) \cdot q(x)) \cdot r(x) = \sum_t e_t x^t$ where

$$e_t = \sum_{s+k=t} d_s c_k = \sum_{s+k=t} \left(\sum_{i+j=s} a_i b_j\right) c_k = \sum_{i+j+k=t} a_i b_j c_k.$$

It can similarly be proved that $p(x) \cdot (q(x) \cdot r(x))$ has the same coefficients. □

The next theorem describes a property of the functions defined by polynomials.

Theorem. *Suppose $q(x), r(x) \in R[x]$ and let $p(x) = q(x) \cdot r(x)$. If R is commutative, then $p(c) = q(c) \cdot r(c)$ for all $c \in R$.*

Proof sketch. Write $q(x) = \sum_i a_i x^i$ and $r(x) = \sum_j b_j x^j$. Then

$$\begin{aligned}q(c) \cdot r(c) &= (a_0 + a_1 c + a_2 c^2 + \cdots) \cdot (b_0 + b_1 c + c_2 c^2 + \cdots) \\ &= a_0 b_0 + a_0(b_1 c) + a_0(b_2 c^2) + a_0(b_3 c^3) + \cdots \\ &\qquad + (a_1 c) b_0 + (a_1 c)(b_1 c) + (a_1 c)(b_2 c^2) + \cdots \\ &\qquad\qquad + (a_2 c^2) b_0 + (a_2 c^2)(b_1 c) + \cdots\end{aligned}$$

If R is commutative, then the terms can be rearranged to get

$$a_0 b_0 + (a_0 b_1 + a_1 b_0)c + (a_0 b_2 + a_1 b_1 + a_2 b_0)c^2 + \cdots = p(c).$$ □

We can generalize this as follows. Given a ring R, its **center** is the set $Z(R) = \{a \in R : ab = ba \text{ for all } b \in R\}$. $Z(R)$ is a subring of R.

Corollary. *Suppose $q(x), r(x) \in R[x]$ and let $p(x) = q(x) \cdot r(x)$. then $p(c) = q(c) \cdot r(c)$ for all $c \in Z(R)$.*

21. Oct 28 – Homomorphisms, ideals

Section 7.3

Definition. Let R, S be rings. A function $\varphi : R \to S$ is a **homomorphism** (of rings) if

(1) $\varphi(a+b) = \varphi(a) + \varphi(b)$ for all $a, b \in R$.
(2) $\varphi(ab) = \varphi(a)\varphi(b)$ for all $a, b \in R$.
(3) $\varphi(1_R) = 1_S$.

Examples.

(1) $\mathbb{Z} \to \mathbb{Z}_n$, $k \mapsto k \pmod n$.
(2) If R is a commutative ring and $c \in R$, then $\varphi_c : R[x] \to R$ given by $\varphi_c(p(x)) = p(c)$. Called "evaluation at c."

    We saw yesterday that φ_c preserves multiplication. Preserving addition is easy. If $p(x)$ is 1, then $p(c)$ is also 1, so $\varphi_c(1) = 1$.

Suppose $\varphi : R \to S$ is a ring homomorphism. Then it is automatically a group homomorphism $\varphi : (R, +) \to (S, +)$. Hence it has a kernel,

$$\ker(\varphi) = \{a \in R \,:\, \varphi(a) = 0_S\}.$$

Furthermore, φ is injective iff $\ker(\varphi) = \{0_R\}$.

Definition. As in the case of groups,

(1) An **isomorphism** is a bijective homomorphism.
(2) Write $R \cong S$ if there exists an isomorphism from R to S.

Definition. Let R be a ring and $I \subseteq R$.

(1) I is a **left ideal** of R if
    (a) I is a subgroup of $(R, +)$.
    (b) If $r \in R$ and $a \in I$, then $ra \in I$.
(2) Right ideals are defined dually ($a \in I, r \in R \implies ar \in I$).
(3) I is an **ideal** if it is both a left and right ideal.

Warning: Dummit & Foote say that an ideal must also be a subring of R. They mean "subrng."

Proposition. *If I is an ideal of R and $1 \in I$, then $I = R$.*

Proof. For every every $r \in R$ we have $r \in R, 1 \in I \implies r1 = r \in I$, so $R \subseteq I$, so $R = I$. □

Proposition 7.5. *Let R, S be rings and $\varphi : R \to S$ a homomorphism.*

(1) $\mathrm{Im}(\varphi)$ *is a subring of S.*
(2) $\ker(\varphi)$ *is an ideal of R.*

Proof. (1) is routine. Focus on (2). We already know that $\ker(\varphi)$ is a (normal) subgroup of $(R,+)$. Suppose $a \in \ker(\varphi)$ and $r \in R$. Then $\varphi(ra) = \varphi(r)\varphi(a) = \varphi(r) \cdot 0_S = 0_S$, proving $ra \in \ker(\varphi)$. A similar proof shows $ar \in \ker(\varphi)$. □

Suppose I is an ideal of R. Then I is a (normal) subgroup of the group $(R,+)$, so we may form the quotient group $(R,+)/I$. Its elements are the cosets of I, which we write additively as $a + I$. The group operation is $(a + I) + (b + I) = (a + b) + I$.

Claim. *The rule* $(a + I) \cdot (b + I) := (ab) + I$ *defines an operation* $\cdot$ *on* R/I.

Proof. We must show that the rule is well-defined. Suppose $a + I = a' + I$ and $b + I = b' + I$, so $a - a' \in I$ and $b - b' \in I$. We must show $(ab) + I = (a'b') + I$, and to do that it suffices to show $ab - a'b' \in I$. Well

$$\begin{aligned} ab - a'b' &= ab - a'b + a'b - a'b' \\ &= (a - a')b + a'(b - b'). \end{aligned}$$

Since I is an ideal and $a - a', b - b' \in I$, the above expression is in I as required. □

Claim. *If* R *is a ring and* I *is an ideal, then* $(R/I, +, \cdot)$ *is a ring.*

We call $(R/I, +, \cdot)$ a **quotient ring** and denote it R/I.

Theorem 7.7 (First Isomorphism Theorem for rings)**.** *Suppose* R, S *are rings and* $\varphi : R \to S$ *is a surjective homomorphism. Then* $R/\ker(\varphi) \cong S$.

An important example of a subring of R is $\mathbb{Z}1 = \{n1 : n \in \mathbb{Z}\}$. It is certainly a subgroup of $(R,+)$ and contains 1. Must check closure under products. Given $m1, n1 \in \mathbb{Z}1$, assume first that $m, n > 0$. Thus

$$m1 \cdot n1 = (\underbrace{1 + 1 + \cdots + 1}_{m}) \cdot (\underbrace{1 + 1 + \cdots + 1}_{n}).$$

Using distributivity repeatedly, one can prove that this equals $(mn)1$ so is in $\mathbb{Z}1$. (One also must check the cases where $m < 0$ or $n < 0$.) We call $\mathbb{Z}1$ the *prime subring* of R and denote it R_0. It is the smallest subring of R, contained in all other subrings.

22. Oct 30 – Principal ideals

Let R be a ring. Recall that $R_0 = \mathbb{Z}1 = \{n1 : n \in \mathbb{Z}\}$. The same calculations that showed that R_0 is a subring of R also show that the function $\varphi : \mathbb{Z} \to R_0$ given by $\varphi(n) = n1$ is a ring homomorphism. It is obviously surjective. By the First Isomorphism Theorem, $\mathbb{Z}/\ker(\varphi) \cong R_0$. Since $\ker(\varphi)$ is a subgroup of $(\mathbb{Z}, +)$, it must equal $n\mathbb{Z}$ for some $n \geq 0$. If $n = 0$ then φ is injective, so is an isomorphism from $\mathbb{Z}$ to R_0. If $n > 0$ then $\mathbb{Z}/n\mathbb{Z} \cong R_0$.

Definition. Let R be a ring. The **characteristic** of R is the integer n in the previous discussion.

Section 7.4.

Definition. Let R be a ring and $a \in R$.

(1) $Ra = \{ra : r \in R\}$.
(2) $aR = \{ar : r \in R\}$.
(3) (a) denotes the smallest ideal of R containing a. (More precisely, (a) is the intersection of all ideals containing a.)

We call (a) the **principal ideal generated by** a.

Lemma. *Suppose R is a ring and $a \in R$.*

(1) *Ra is a left ideal. It is the smallest left ideal of R containing a.*
(2) *Similarly, aR is the smallest right ideal of R containing a.*
(3) *$Ra \cup aR \subseteq (a)$.*

Proof. (1) Obviously $Ra \neq \varnothing$. Suppose $ra, sa \in Ra$. Then $ra + sa = (r + s)a \in Ra$ and $-(ra) = (-r)a \in Ra$, so Ra is a subgroup of $(R, +)$. Clearly if $ra \in Ra$ and $s \in R$ then $s(ra) = (sr)a \in Ra$, proving Ra is a left ideal. Clearly $a = 1a$ so $a \in Ra$. Now suppose that I is any left ideal of R containing a. Since $a \in I$ and I is a left ideal, we get $ra \in I$ for all $r \in R$; hence $Ra \subseteq I$. This proves Ra is contained in every left ideal containing a, so is the smallest such left ideal.

(2) is proved similarly. (3) Since (a) is an ideal and $a \in (a)$, it follows that $ra, ar \in (a)$ for all $r \in R$. This proves $Ra \cup aR \subseteq (a)$. □

Note: If R is commutative, then $Ra = aR$ and Ra is an ideal of R containing a. Since (a) is by definition the smallest ideal containing a, we get $(a) \subseteq Ra$. We already know that $Ra \subseteq (a)$. Hence $\boxed{(a) = Ra = aR \text{ if } R \text{ is commutative}}$.

Example. Consider the ring $\mathbb{R}[x]$. Let $I = (x^2 + 1)$, the principal ideal generated by $x^2 + 1$. Thus

$$I = \{(x^2 + 1)q(x) : q(x) \in \mathbb{R}[x]\} = \{f(x) \in \mathbb{R}[x] : x^2 + 1 \text{ is a factor of } f(x)\}.$$

(In the expression $(x^2+1)q(x)$, (x^2+1) does NOT denote the ideal I; the parentheses are just being used to surround the factor of $x^2 + 1$. It will be your job to recognize

when parentheses are being used as brackets and when they are being used to name a principal ideal.)

I is an ideal, so we can form the quotient ring $\mathbb{R}[x]/I$. What is this quotient ring isomorphic to? Take an arbitrary element, i.e. a coset $f(x)+I$. Divide $f(x)$ by x^2+1 to get

$$f(x) = (x^2+1)q(x) + (a+bx),$$

Hence

$$\begin{aligned} f(x)+I &= [(x^2+1)q(x)+I] + [(a+bx)+I] \\ &= I + [(a+bx)+I] \qquad \text{because } x^2+1 \in I \\ &= (a+bx)+I \qquad \text{because } I \text{ is the zero element of } \mathbb{R}[x]/I. \end{aligned}$$

In other words, every coset of I can be expressed as $(a+bx)+I$ for some $a, b \in \mathbb{R}$. Hence

$$\mathbb{R}[x]/I = \{(a+bx)+I \,:\, a,b \in \mathbb{R}\}.$$

Let's explore how addition and multiplication work in $\mathbb{R}[x]/I$. Let $(a+bx)+I$, $(c+dx)+I$ be two elements of $R[x]/I$. Their sum is easily computed.

$$\begin{aligned} [(a+bx)+I] + [(c+dx)+I] &= [(a+bx)+(c+dx)]+I \\ &= [(a+c)+(b+d)x]+I. \end{aligned}$$

Similarly,

$$\begin{aligned} [(a+bx)+I]\cdot[(c+dx)+I] &= [(a+bx)\cdot(c+dx)]+I \\ &= [(ac)+(ad+bc)x+(bd)x^2]+I. \end{aligned}$$

We can simplify this last expression as follows. Since $x^2+1 \in I$ we get $x^2+I = -1+I$, so $(bd)x^2+I = -bd+I$, so

$$\begin{aligned} [(a+bx)+I]\cdot[(c+dx)+I] &= [(ac)+(ad+bc)x-bd]+I \\ &= [(ac-bd)+(ad+bc)x]+I. \end{aligned}$$

This resembles multiplication in $\mathbb{C}$. We might conjecture that $\mathbb{R}[x]/I \cong \mathbb{C}$. To prove this conjecture, define $\varphi : \mathbb{R}[x] \to \mathbb{C}$ by $\varphi(p(x)) = p(i)$. One can show that φ is a homomorphism. (See the 2nd example from Oct 28.) Obviously φ is surjective, since for any complex number $a+ib$ we have $a+ib = \varphi(a+bx)$. I claim that $\ker(\varphi) = I$. Indeed, if $f(x) \in \ker(\varphi)$, i.e., $f(i) = 0$, then both $i, -i$ are roots of $f(x)$ so x^2+1 is a factor of $f(x)$, meaning $f(x) \in (x^2+1) = I$. This proves $\ker(f) \subseteq I$. For the converse inclusion, note that $\ker(\varphi)$ is an ideal which contains x^2+1 (obviously), so $(x^2+1) \subseteq \ker(\varphi)$ (since (x^2+1) is contained in every ideal containing x^2+1). This proves $\ker(\varphi) = I$.

Now apply the First Isomorphism Theorem to get $\mathbb{R}[x]/I \cong \mathbb{C}$.

23. Nov 3 – Maximal ideals

The ideals of a ring R are ordered by inclusion and hence form a partially ordered set (poset). We can schematically draw this poset with R at the top, $\{0\}$ at the bottom, and other ideals in between.

Lemma. *Suppose I, J are ideals of R.*

(1) *$I \cap J$ is an ideal; it is the largest ideal of R contained in both I and J.*
(2) *$I + J := \{a + b : a \in I \text{ and } b \in J\}$ is the smallest ideal of R containing both I and J.*

Proof. (2) I, J are both (normal) subgroups of $(R, +)$, so $I+J$ is also a subgroup and containing I and J. Suppose $a+b \in I+J$ and $r \in R$. Then $r(a+b) = ra+rb \in I+J$ and similarly $(a+b)r = ar+br \in I+J$, so $I+J$ is an ideal. We've already noted $I, J \subseteq I+J$. Suppose K is any other ideal with $I, J \subseteq K$. Then for all $a+b \in I+J$ we have $a, b \in K$ so $a+b \in K$, proving $I+J \subseteq K$. □

Definition. Let R be a ring.

(1) An ideal I is **proper** if $I \neq R$. (Equivalently, if $1 \notin I$.)
(2) If I, J are ideals, then J **properly contains** I if $I \subseteq J$ and $I \neq J$.
(3) I is a **maximal ideal** if (i) it is a proper ideal, and (ii) the only ideal properly containing it is R.

Proposition 7.12. *Suppose R is a commutative ring and I is an ideal. R/I is a field iff I is a maximal ideal.*

Proof. Throughout the proof, if $a \in R$ then $\overline{a}$ denotes $a + I \in R/I$. In particular, $\overline{0} = 0 + I$ is the zero of R/I and $\overline{1} = 1 + I$ is the multiplicative identity of R/I.

($\Rightarrow$) Assume R/I is a field. Then $\overline{0} \neq \overline{1}$, meaning $I \neq 1+I$, so $1 \notin I$, so I is proper. Suppose J is an ideal properly containing I. Pick $a \in J \backslash I$. Thus $a+I \neq I$, i.e., $\overline{a} \neq \overline{0}$. As R/I is a field, there exists $\overline{b} \in R/I$ such that $\overline{a} \cdot \overline{b} = \overline{1}$, i.e., $(a+I)(b+I) = 1+I$, so $1 = ab + c$ for some $c \in I$. As $a, c \in J$ and J is an ideal, we get $1 \in J$ so $J = R$. This proves I is maximal.

($\Leftarrow$) Suppose I is maximal. We run through the defining properties of being a field.

(1) R/I is commutative (because R is).
(2) $1 \notin I$ because I is proper, so $I \neq 1 + I$, so $\overline{0} \neq \overline{1}$.
(3) Let $\overline{a} \in R/I$ with $\overline{a} \neq \overline{0}$. (Thus $a \notin I$.) We must show that $\overline{a}$ has a multiplicative inverse in R/I. Recall that $(a) = Ra$. By hypothesis, $a \notin I$, but clearly $a \in (a) + I$, so $(a) + I$ properly contains I, so $(a) + I = R$. In particular, $1 \in (a) + I$, so there exists $r \in R$ and $c \in I$ such that $1 = ar + c$. Hence $1+I = ar+I = (a+I)(r+I)$, meaning $\overline{1} = \overline{a} \cdot \overline{r}$, so $\overline{r}$ is a multiplicative inverse to $\overline{a}$. □

24. Nov 4 – Prime ideals, Zorn's Lemma

Recall that a ring is an integral domain if it is commutative, satisfies $0 \neq 1$, and has no zero divisors. This last condition is equivalent to $ab = 0$ implying $a = 0$ or $b = 0$.

Suppose R is commutative and I is an ideal. What properties of I determine whether R/I is an integral domain? R/I is already commutative (because R is). Clearly we need I to be proper (so $\overline{0} \neq \overline{1}$). To achieve the final condition, we need $\overline{a} \cdot \overline{b} = \overline{0}$ to imply $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$.

- $\overline{a} \cdot \overline{b} = \overline{0}$ means $(a + I)(b + I) = I$, i.e., $ab \in I$.
- $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$ means $a + I = I$ or $b + I = I$, i.e., $a \in I$ or $b \in I$.

Thus we need: $ab \in I$ implies $a \in I$ or $b \in I$. Ideals with this property (and proper) are called **prime ideals**.

Every maximal ideal (of a commutative ring) is a prime ideal (because R/I is a field, so is an integral domain). The converse is not true. (Example: in $\mathbb{Z}$, $\{0\}$ is a prime ideal but is clearly not a maximal ideal.)

Proposition 7.11. *Let R be a ring. Every proper ideal of R is contained in a maximal ideal of R.*

Proof. Here is the idea of the proof. Let I be a proper ideal of R. define $I_0 = I$. If I_0 is maximal then we're done. Otherwise, there exists a proper ideal I_1 properly containing I_0. If I_1 is maximal, we're done, and if not, then there exists a proper ideal I_2 properly containing I_1. In this way we either reach a maximal ideal or we construct an infinite sequence of proper ideals:

$$I = I_0 \subset I_1 \subset I_2 \subset \cdots \subset I_n \subset \cdots$$

Define $I_\infty = \bigcup_{n=0}^{\infty} I_n$. We claim that I_∞ is a proper ideal of R.

- Suppose $a, b \in I_\infty$. Then there exists n with $a, b \in I_n$. so $a+b, -a \in I_n \subseteq I_\infty$.
- Suppose $a \in I_\infty$ and $r \in R$. Then $a \in I_n$ for some n. Hence $ra \in I_n \subseteq I_\infty$.
- Thus I_∞ is an ideal. To show it is proper, suppose instead that $I_\infty = R$. Then $1 \in I_\infty$. Hence $1 \in I_n$ for some n. But then I_n isn't proper, contradiction. Hence I_∞ is proper.

We continue the argument. If I_∞ is maximal we're done. Otherwise, there exists a proper ideal $I_{\infty+1}$ properly containing I_∞. Continue: either a maximal ideal $I_{\infty+n}$ is found, or we get another infinite sequence of proper ideals:

$$I_\infty \subset I_{\infty+1} \subset I_{\infty+2} \subset \cdots \subset I_{\infty+n} \subset \cdots$$

Define $I_{\infty+\infty} = \bigcup_{n=0}^{\infty} I_{\infty+n}$. Again this is a proper ideal.

The intuition is that this cannot go on forever. To prove it, we must clarify what we mean by "forever." This is the job of set theory; for example, countable sequences

(no matter how many times applied) do not capture "forever." We can "sweep this under the carpet" by a trick from set theory.

Definition. A **chain of proper ideals** is set S of proper ideals with the property that for all $I, J \in S$, either $I \subseteq J$ or $J \subseteq I$. (Note: S can be uncountable.)

By a similar argument as above, if S is a chain of proper ideals, then $\bigcup_{I \in S} I$ is still a proper ideal. In particular, for every chain of proper ideals there exists a proper ideal containing all the elements of the chain.

Now let $\mathcal{I}(R)$ be the set of all proper ideals of R. The relation $\subseteq$ is a partial ordering of $\mathcal{I}(R)$ (reflexive, antisymmetric, transitive). We have proved that every chain in $(\mathcal{I}(R), \subseteq)$ has an upper bound in $\mathcal{I}(R)$.

Lemma (Zorn's Lemma)**.** *Suppose $(A, \leq)$ is a set equipped with a partial order. If every chain in $(A, \leq)$ has an upper bound in A, then every element of A lies below a maximal element.*

If we apply this with $(A, \leq) = (\mathcal{I}(R), \subseteq)$ we get the Proposition. □

Commentary. The proof of Zorn's Lemma is a souped-up version of the intuitive proof presented above. It constructs a "transfinite" chain

$$a = a_0 < a_1 < a_2 < \cdots < a_\infty < a_{\infty+1} < \cdots < a_{\infty+\infty} < \cdots$$

of elements of A. However, "constructs" is not quite right. At stage α, we have an element a_α which is not maximal. To "construct" $a_{\alpha+1}$, we need to **choose** one element (from potentially many) which properly extends a_α. There may be no natural way to do this (even though we know some such element must exist). Some mathematicians and philosophers have objected to "constructions" that require infinitely many ad hoc choices. The Axiom of Choice (in set theory) asserts that constructions of this kind are OK, so Zorn's Lemma is correct (unless the Axiom of Choice is false . . .).

25. Nov 6 – Rings of fractions

Section 7.5. Suppose R is an integral domain and $D \subseteq R$ is a subset of R satisfying

(1) $1 \in D$.
(2) $0 \notin D$.
(3) D is closed under multiplication (i.e., $a, b \in D$ implies $ab \in D$).

(For example, the set $D = R \setminus \{0\}$ satisfies these properties.)

I will show that the standard construction of $\mathbb{Q}$ (as fractions n/d where $n, d \in \mathbb{Z}$ with $d \neq 0$) can be carried out to construct an integral domain of "fractions" r/d where $r \in R$ and $d \in D$.

Let $\mathcal{F} = R \times D = \{(r, d) : r \in R, d \in D\}$. Define a relation $\sim$ on $\mathcal{F}$ by

$$(r, d) \sim (s, e) \qquad \text{iff} \qquad re = sd.$$

Claim. *$\sim$ is an equivalence relation on $\mathcal{F}$.*

Proof. It is easily shown to be reflexive and symmetric. For transitivity, suppose $(r, d) \sim (s, e)$ and $(s, e) \sim (t, f)$. Thus $re = sd$ and $sf = te$. Hence

$$ref = sdf = sfd = ted,$$

so $(rf - td)e = 0$. As R is an integral domain, we can deduce $rf - td = 0$ or $e = 0$. However, $e \in D$ so $e \neq 0$ by (2). Hence $rf - td = 0$, so $(r, d) \sim (t, f)$, proving $\sim$ is transitive. □

For $(r, d) \in \mathcal{F}$ let r/d denote the equivalence class of $\sim$ containing (r, d). That is,

$$r/d = \{(s, e) \in \mathcal{F} : (r, d) \sim (s, e)\}.$$

Define F to be the set of these equivalence classes:

$$F = \{r/d : (r, d) \in \mathcal{F}\}.$$

Note that $r/d = s/e$ <u>means</u> $(r, d) \sim (s, e)$.

By (1), $r/1 \in F$ for all $r \in R$. We call $r/1$ the **image** of r. Note that distinct elements of R have distinct images in F, since $r/1 = s/1$ implies $(r, 1) \sim (s, 1)$, i.e., $r1 = s1$, i.e., $r = s$.

Next, we define $+$ and $\cdot$ on F in the "grade school" way:

$$\begin{aligned} r/d + s/e &:= (re + sd)/de \\ (r/d) \cdot (s/e) &:= rs/de. \end{aligned}$$

Note that $d, e \in D$ implies $de \in D$ by (3), so the right-hand sides make sense.

Theorem.

(1) *$+$ and $\cdot$ are well-defined.*
(2) *$(F, +, \cdot)$ is an integral domain.*
(3) *$\{r/1 : r \in R\}$ is a subring of F isomorphic to R.*

Commentary. The assertion that "+ is well-defined" means the following: for all $r_1, r_2, s_1, s_2 \in R$ and all $d_1, d_2, e_1, e_2 \in D$, if $r_1/d_1 = r_2/d_2$ and $s_1/e_1 = s_2/e_2$, then $(r_1e_1 + s_1d_1)/d_1e_1 = (r_2e_2 + s_2d_2)/d_2e_2$; equivalently, if $(r_1, d_1) \sim (r_2, d_2)$ and $(s_1, e_1) \sim (s_2, e_2)$, then $((r_1e_1 + s_1d_1), d_1e_1) \sim ((r_2e_2 + s_2d_2), d_2e_2)$. The proof of this claim, and everything else claimed in this theorem, is left as an excellent exercise. (In particular, $0/1$ will be the zero element and $1/1$ will be the identity element of F.)

In practice, we identify each element $r \in R$ with its image $r/1 \in F$. This makes R a virtual subring of F.

Definition. The ring F constructed above is called the **ring of fractions of R over D**, and is denoted $D^{-1}R$.

Claim. *If $D = R \setminus \{0\}$, then F is a field.*

Proof. We already know that $D^{-1}R$ is an integral domain, so it remains to show that every nonzero element has a multiplicative inverse. Suppose $r/d \in D^{-1}R$ is nonzero, i.e., $r/d \neq 0/1$. This means $r/d \neq 0/1$, i.e., $(r, d) \not\sim (0, 1)$, i.e., $r1 \neq 0d$, so $r \neq 0$. So $r \in D$. So $d/r \in F$, and clearly $(r/d) \cdot (d/r) = (rd, rd) = 1/1$. Hence r/d is invertible with inverse d/r. □

Example. Let $R = \mathbb{R}[x]$ and $D = R \setminus \{0\}$. Then $D^{-1}R$ is a field containing $\mathbb{R}[x]$ (virtually) as a subring, and every element of D^{-1}/R can be expressed as a fraction $p(x)/q(x)$ for some $p(x), q(x) \in \mathbb{R}[x]$ with $q(x) \neq 0$. This field is denoted $\mathbb{R}(x)$ and is called the field of **rational functions over** $\mathbb{R}$, but note that the elements of $\mathbb{R}(x)$ are **not** functions; they are equivalence classes of a relation $\sim$ defined on the set $\mathcal{F}$ of pairs $(p(x), q(x))$.

Example. Let $R = \mathbb{Z}$ and $D = \{d \in \mathbb{Z} : 3 \not| d\}$. D satisfies assumptions (1)–(3), so the above construction gives an integral domain $D^{-1}\mathbb{Z}$ properly containing $\mathbb{Z}$, in which every integer in D becomes a unit. More precisely,

$$D^{-1}\mathbb{Z} = \{n/d : n, d \in \mathbb{Z},\, d \not\equiv 0 \pmod 3\}.$$

Note that $D^{-1}\mathbb{Z}$ is not a field, since e.g. the element 3 is not invertible.

Example. More generally, suppose R is an integral domain and I is a prime ideal of R. Let $D = R \setminus I$, i.e., the complement of I. Then D satisfies assumptions (1)–(3) (exercise), so $D^{-1}R$ is defined. It is called the **localization of R at the prime ideal I**.

26. Nov 10 – Chinese Remainder Theorem

Section 7.6

(Took up problem 2(a) from Assignment 7.)

Consider the ring $\mathbb{Z}$. Fix $m \geq 1$, let $I = (m)$, and consider the quotient ring $\mathbb{Z}/I$. Note that for all $a, b \in \mathbb{Z}$,

$$\begin{aligned} a + I = b + I &\iff b - a \in I \\ &\iff b - a \in (m) \\ &\iff m | b - a \\ &\iff a \equiv b \pmod{m}. \end{aligned}$$

This motivates the next definition.

Definition. If R is a ring, I is an ideal, and $a, b \in R$, then we write

$$a \equiv b \pmod{I}$$

to mean $a + I = b + I$ (equivalently, $b - a \in I$).

Consider again the ring $\mathbb{Z}$. Suppose $m, n \in \mathbb{Z}$ are **coprime**, i.e., $\gcd(m, n) = 1$. We know from MATH 135 that there exist $r, s \in \mathbb{Z}$ with $rm + sn = 1$.

Now let $I = (m)$ and $J = (n)$. By the above, we have $rm \in I$ and $sn \in J$, so $1 = rm + sn \in I + J$. Since $I + J$ is an ideal, this proves $I + J = \mathbb{Z}$. This motivates:

Definition. Let R be a ring. Two ideals I, J are **comaximal** if $I + J = R$.

Theorem (Chinese Remainder Theorem)**.** *Suppose R is a ring and I, J are comaximal ideals. For all $a, b \in R$ then there exists $c \in R$ such that*

$$\begin{aligned} c &\equiv a \pmod{I}, \quad \textit{and} \\ c &\equiv b \pmod{J}. \end{aligned}$$

Proof. Because $I + J = R$, there exist $e \in I$ and $f \in J$ with $1 = e + f$. Let $c = af + be$. Observe that

$$\begin{aligned} e &\equiv 0 \pmod{I} \qquad \text{as } e \in I \\ f &\equiv 1 \pmod{I} \qquad \text{as } 1 - f = e \in I \end{aligned}$$

Hence

$$c = af + be \equiv a1 + b0 \pmod{I}$$

i.e., $c \equiv 0 \pmod{I}$. A similar proof shows $c \equiv b \pmod{J}$. □

Definition. Suppose $R = (R, +, \cdot)$ and $S = (S, +, \cdot)$ are rings. Their **direct product** is $(R \times S, +, \cdot)$ where $+$ and $\cdot$ are defined coordinatewise:

$$\begin{aligned} (r_1, s_1) + (r_2, s_2) &:= (r_1 + r_2,\ s_1 + s_2) \\ (r_1, s_1) \cdot (r_2, s_2) &:= (r_1 \cdot r_2,\ s_1 \cdot s_2). \end{aligned}$$

It is a ring. $(R \times S, +)$ is just the direct product of the groups $(R, +)$ and $(S, +)$. The zero element of $R \times S$ is $(0_R, 0_S)$. The identity element of $R \times S$ is $(1_R, 1_S)$.

On Sept 23 I explained the test for recognizing direct products of groups: if $H, K \lhd G$ and $H \cap K = \{1\}$ and $HK = G$, then $G \cong H \times K$. On Oct 2 I showed that, under the same hypotheses, we have $G/H \cong K$ and $G/K \cong H$. Hence

$$G \cong G/H \times G/K.$$

This last fact has a version that works for rings.

Corollary. *Suppose R is a ring and I, J are comaximal ideals.*

(1) $R/(I \cap J) \cong R/I \times R/J$.
(2) *If $I \cap J = \{0\}$ then $R \cong R/I \times R/J$.*

[Proof to follow tomorrow]

27. Nov 11 – PIDs

Proof of yesterday's Corollary. (2) follows from (1) since $R \cong R/\{0\}$. To prove (1), define $\varphi : R \to R/I \times R/J$ by

$$\varphi(r) = (r + I,\, r + J).$$

The idea is to show that φ is a surjective ring homomorphism and apply the 1st Isomorphism Theorem. I won't check that φ is a homomorphism (but it is a good exercise in understanding definitions).

I will prove that φ is surjective. Suppose $(a + I,\, b + J)$ is an arbitrary element of $R/I \times R/J$. By the Chinese Remainder Theorem, there exists $c \in R$ with

$$\begin{aligned} c &\equiv a \pmod{I}, \quad \text{and} \\ c &\equiv b \pmod{J}. \end{aligned}$$

Thus

$$\varphi(c) = (c + I,\, c + J) = (a + I,\, b + J).$$

Finally, we compute the $\ker(\varphi)$. If $r \in R$, then

$$\begin{aligned} r \in \ker(\varphi) &\iff \varphi(r) = 0 = (I, J) \\ &\iff (r + I,\, r + J) = (I, J) \\ &\iff r \in I \text{ and } r \in J \\ &\iff r \in I \cap J. \end{aligned}$$

Hence $\ker(\varphi) = I \cap J$. □

Example. Let $R = \mathbb{Z}$ and $I = (m)$ and $J = (n)$ where $\gcd(m, n) = 1$. Then

$$\begin{aligned} I \cap J &= \{a \in \mathbb{Z} : m|a \text{ and } n|a\} \\ &= \{a \in \mathbb{Z} : mn|a\} \quad (\text{because } \gcd(m, n) = 1 \\ &= (mn). \end{aligned}$$

Thus $\mathbb{Z}/I \cong \mathbb{Z}_m$, $\mathbb{Z}/J \cong \mathbb{Z}_n$, and $\mathbb{Z}/(I \cap J) \cong \mathbb{Z}_{mn}$, so the CRT gives

$$\mathbb{Z}_{mn} \cong \mathbb{Z}_m \times \mathbb{Z}_n.$$

Chapter 8 – Principal ideal domains

Proposition. *Every ideal of $\mathbb{Z}$ is principal.*

Proof. Suppose I is an ideal of $\mathbb{Z}$. If $I = \{0\}$ then $I = (0)$. Otherwise, pick $a \in I$ with $a \neq 0$ and $|a|$ minimum. Clearly $(a) \subseteq I$. To prove $\supseteq$, assume $b \in I$. Divide b by a to get quotient q and remainder r, so

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

$a, b \in I$ implies $r = b - aq \in I$. Hence $r = 0$, so $b = aq$, proving $b \in (a)$. □

Definition. An ring R is a **Principal Ideal Domain** (or PID) if

(1) R is an integral domain (commutative, $0 \neq 1$, no zero divisors).
(2) Every ideal of R is principal.

Example. The following are examples of PIDs.

(1) $\mathbb{Z}$.
(2) Any field. (Because a field F only has two ideals: $\{0\} = (0)$ and $F = (1)$.)
(3) $\mathbb{R}[x]$.

Proof. It is an integral domain. Let I be an ideal. If $I = \{0\}$ then $I = (0)$. Otherwise pick $f(x) \in I$ with $f(x) \neq 0$ and with $\deg(f(x))$ minimum. Clearly $(f(x)) \subseteq I$. For $\supseteq$, assume $g(x) \in I$. Divide $g(x)$ by $f(x)$ to get quotient $q(x)$ and remainder $r(x)$ (in $\mathbb{R}[x]$), so

$$g(x) = f(x)q(x) + r(x), \qquad r(x) = 0 \text{ or } \deg(r(x)) < \deg(f(x)).$$

$f(x), g(x) \in I$ implies $r(x) = g(x) - f(x)q(x) \in I$. Hence $r(x) = 0$, so $g(x) = f(x)q(x)$, proving $g(x) \in (f(x))$. □

(4) More generally, $F[x]$ where F is a field. (Same argument, using the division algorithm in $F[x]$.)
(5) Even more generally, any integral domain for which we have a "division algorithm" which, given any $a, b \in R$ with $a \neq 0$, produces a quotient/remainder pair $q, r \in R$ satisfying
   - $b = aq + r$.
   - r is "strictly simpler" than a.

   There are several ways to formulate this. A standard way leads to the definition of *Euclidean domains*. $\mathbb{Z}[x]$ and polynomial rings $F[x]$ (where F is a field) are examples of Euclidean domains. We won't study Euclidean domains in this course.

28. Nov 13 – Primes and irreducibles

Recall that if R is a ring than $R^\times$ denotes the set of **units** (invertible elements) of R. Also recall (Assign. 7) that in a commutative ring R we say a **divides** b and write $a|b$ if $b = ar$ for some $r \in R$.

Example.

(1) $\mathbb{Z}^\times = \{1, -1\}$.
(2) $\mathbb{Z}[i] = \{a + ib : a, b \in \mathbb{Z}\}$ is the ring of **Gaussian integers**. It is a subring of $\mathbb{C}$, so is an integral domain. Its set of units is $\mathbb{Z}[i] = \{1, i, -1, -i\}$.
(3) $\mathbb{Z}[\sqrt{3}] = \{a + b\sqrt{3} : a, b \in \mathbb{Z}\}$ is a subring of $\mathbb{R}$. Its set of units is infinite, containing e.g. $\pm 1,\ \pm 2 \pm \sqrt{3},\ \pm 7 \pm 4\sqrt{3}, \ldots$.
(4) $\mathbb{Z}[\sqrt{-5}] = \{a + ib\sqrt{5} : a, b \in \mathbb{Z}\}$ is a subring of $\mathbb{C}$. Its set of units is $\{1, -1\}$.

Lemma. *In a commutative ring R, an element u is a unit iff $u|1$.*

Proof. $u|1$ iff $1 = uv$ for some $v \in R$, iff $v = u^{-1}$, i.e., $u \in R^\times$. □

Corollary. *In a commutative ring R, u is a unit iff $(u) = (1)$.*

Proof. $(u) = (1)$ iff $(1) \subseteq (u)$ (since the opposite inclusion is always true, as $(1) = R$). $(1) \subseteq (u)$ iff $u|1$ (by Assign. 7 problem 3(a)). □

Definition. We say that a and b are **associates** and write $a \sim b$ if $a = ub$ for some unit $u \in R^\times$.

Example.

(1) In $\mathbb{Z}$, $a \sim b$ iff $a = \pm b$.
(2) In $\mathbb{R}[x]$, $2x + 3 \sim x + \frac{3}{2}$ since $2x + 3 = 2(x + \frac{3}{2})$ and 2 is invertible in $\mathbb{R}[x]$.

Lemma. *In an integral domain R, $a \sim b$ iff $a|b$ and $b|a$.*

Proof. ($\Rightarrow$). Assume $a \sim b$, so $a = ub$ with $u \in R^\times$. Then obviously $b|a$. And $u^{-1}a = b$ with $u^{-1} \in R$, by assumption, so $a|b$.

($\Leftarrow$). Assume $a|b$ and $b|a$. This means $b = ar$ and $a = bs$ for some $r, s \in R$. Hence $a = bs = (ar)s = a(rs)$, so $a(1 - rs) = 0$. As we are in an integral domain, we can deduce $a = 0$ or $1 - rs = 0$.

Case 1. $a = 0$
Then $b = ar$ implies $b = 0$, so we can write e.g. $a = 1b$. 1 is a unit so $a \sim b$.

Case 2. $1 - rs = 0$
Then $rs = sr = 1$, so s is a unit, so $a = sb$ gives $a \sim b$.

Thus $a \sim b$ in either case, proving ($\Leftarrow$). □

Corollary. *In an integral domain R, $a \sim b$ iff $(a) = (b)$.*

Proof. $(a) = (b)$ iff $(a) \subseteq (b)$ and $(b) \subseteq (a)$. By problem 3(a) of Assign. 7, this is equivalent to $b|a$ and $a|b$. □

Definition. Let R be an integral domain. Assume $a \in R$ with $a \neq 0$ and $a \notin R^\times$.

(1) A **nontrivial factorization** of a is an equation $a = bc$ where $b, c \in R$ and neither b nor c is a unit.
(2) a is **reducible** if it has a nontrivial factorization in R.
(3) Otherwise a is **irreducible** (equivalently, $a = bc$ implies b or c is a unit).
(4) We say that a is a **prime** if for all $b, c \in R$, if $a|bc$ then $a|b$ or $a|c$.

Note that these definitions are always **relative to** R. For example,

- 3 is a both prime and irreducible in $\mathbb{Z}$.
- 3 is reducible in $\mathbb{Z}[\sqrt{3}]$, because $3 = (\sqrt{3})(\sqrt{3})$ is a nontrivial factorization.
- 3 is neither reducible nor irreducible in $\mathbb{R}$, because it is a unit there.

Proposition 8.10. *In an integral domain, every prime is irreducible.*

Proof. Suppose p is prime and $p = bc$. We can write $p1 = bc$, so $p|bc$, so by definition of being a prime, $p|b$ or $p|c$.

Case 1: $p|c$.

We also have $c|p$ (from $p = bc$). So $p \sim c$, say $p = uc$ with $u \in R^\times$. Obviously $c \neq 0$ (as $p \neq 0$), so $bc = uc$ implies $b = u$ so $b \in R^\times$.

Case 2. $p|b$.

Then a similar argument shows $p \sim b$ and $c \in R^\times$. Since either Case 1 or 2 holds, we've shown that if $p = bc$ then b or c is a unit. So p has no nontrivial factorization in R, meaning it is irreducible. □

The converse is not always true, as the next example shows.

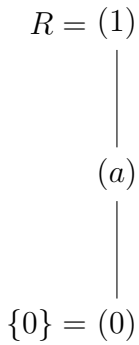
Example. Let R be the set of all complex numbers of the form $a + bi\sqrt{5}$ where $a, b \in \mathbb{Z}$. R is a subring of $\mathbb{C}$ and so is an integral domain. It is possible to show that $R^\times = \{1, -1\}$ and that 3 is irreducible in R, i.e., cannot be factored nontrivially. Let $c = 2 + i\sqrt{5}$ and $d = 2 - i\sqrt{5}$. So $c, d \in R$ and $cd = 4 + 5 = 9$, so $3|cd$. But 3 divides neither c nor d in R (since $\frac{2}{3} \pm \frac{1}{3}i\sqrt{5} \notin R$). Thus 3 is **not** a prime in R.

29. Nov 17 – Complete factorizations

Recap: in an integral domain R, suppose $a \neq 0$ and $a \notin R^\times$.

- A factorization $a = bc$ is **trivial** if b or c is a unit, and is **nontrivial** otherwise.

We can picture nontrivial factorizations in the partially ordered set of ideals of R.

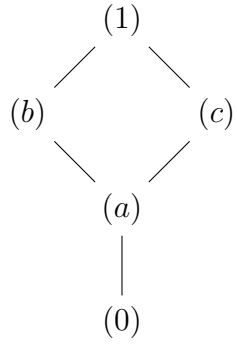


The hypothesis translates as follows:

- $a \neq 0 \iff (a) \neq (0)$.
- $a \notin R^\times \iff (a) \neq (1)$.

Now suppose $a = bc$. Focus on (b).

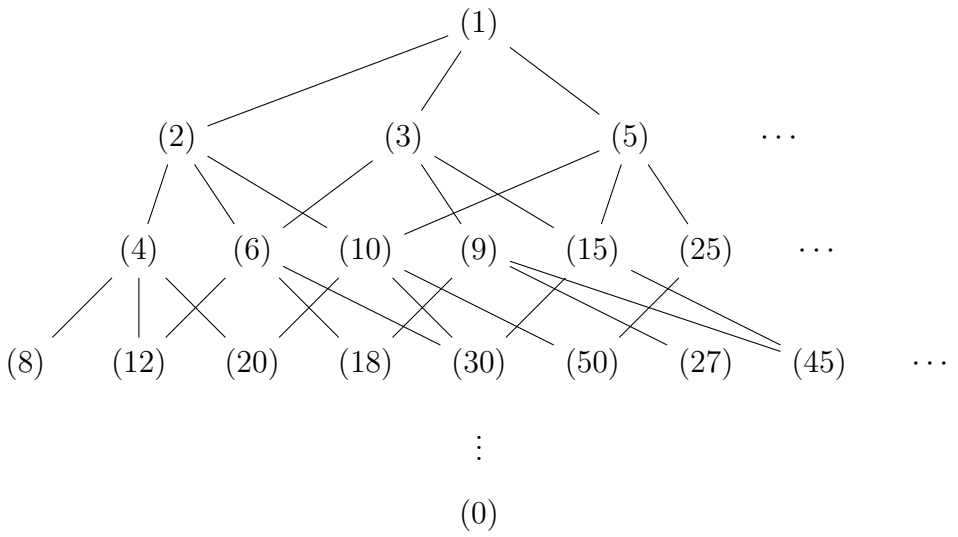
- If this factorization is trivial, then b or c is a unit.
  - If b is a unit, then $(b) = (1)$.
  - If c is a unit, then $b \sim a$, so $(b) = (a)$.
- If the factorization is nontrivial, then neither b nor c is a unit, Because b is not a unit, $(b) \neq (1)$. Because c is not a unit, $b \not\sim a$, so $(b) \neq (a)$. Of course $(b) \subseteq (1)$. Finally, $b|a$ so $(a) \subseteq (b)$. Hence $(a) \subset (b) \subset (1)$. (Similar remarks hold for (c).)



Hence the factorization is nontrivial iff $(a) \subset (b) \subset (1)$. This proves:

Proposition. *Suppose R is an integral domain and $a \in R$. Then a is irreducible iff $(a) \neq (0)$, $(a) \neq (1)$, and there is no principal ideal (b) properly between (a) and (1).*

Example. Draw a picture of the principal ideals in $\mathbb{Z}$.



Notice:

- $6 = 2 \cdot 3$ translates to (2) and (3) above (6), below (1).
- $4 = 2 \cdot 2$ translates to just (2) above (4), below (1).
- Irreducibles (= primes) are just below the top (they are maximal ideals).
- Can also "see" $(12) + (30)$; it must be the smallest ideal containing both (12) and (30). We see that it is (6). Consistent with $6 = \gcd(12, 30)$.

Definition. Suppose R is an integral domain, $a \in R$, $a \neq 0$, and $a \notin R^\times$. A **complete factorization** of a is an equation

$$a = p_1 p_2, \ldots, p_n$$

where $n \geq 1$, $p_1, \ldots, p_n \in R$, and each p_i is irreducible.

Naive algorithm to find a complete factorization. Given $a \in R$ with $a \neq 0$ and $a \notin R^\times$:

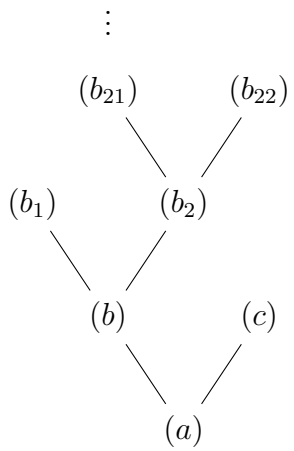
- If a is irreducible, then done.
- Else pick a nontrivial factorization $a = bc$.
- Recursively find complete factorizations for b and c:

$$b = p_1 p_2 \cdots p_n \qquad \text{and} \qquad c = q_1 q_2 \cdots q_m.$$

- Then $a = p_1 p_2 \cdots p_n q_1 q_2 \cdots q_m$ is a complete factorization of a.

There is one potential problem with this algorithm? What is the problem?
(Answer: it may never terminate.)

E.g., b might have a nontrivial factorization $b = b_1 b_2$. Then b_2 might have a nontrivial factorization $b_2 = b_{21} b_{22}$. And so on.



The bad thing (failure to terminate) can **only** happen if there is an **infinite** strictly increasing chain $(a) \subset (b) \subset (b_2) \subset (b_{21}) \subset \cdots$ of principal ideals. This proves:

Proposition. *Suppose R is an integral domain and R does **not** have an infinite strictly increasing chain of principal ideals. Then every $a \in R$ with $a \neq 0$, $a \notin R^\times$ has a complete factorization.*

The previous proposition addresses the existence of complete factorizations. Next we study uniqueness.

Example. In $\mathbb{Z}$, 6 has four complete factorizations:

$$\begin{aligned} 6 &= (2)(3) \\ 6 &= (3)(2) \\ 6 &= (-2)(-3) \\ 6 &= (-3)(-2). \end{aligned}$$

These are "essentially the same" factorization.

Definition. Let R be an integral domain and $a \in R$ with $a \neq 0$, $a \notin R^\times$.

(1) Two complete factorizations of a,

$$a = p_1 p_2 \cdots p_n \qquad \text{and} \qquad a = q_1 q_2 \cdots q_m$$

are **essentially the same up** provided
   (a) $m = n$, and
   (b) After a suitable reordering of the q_i's we have $p_i \sim q_i$ for all $i = 1, \ldots, n$.

(2) We say that **complete factorizations in R are unique, when they exist** provided for any $a \in R$ with $a \neq 0$ and $a \notin R^\times$, **if** a has a complete factorization, **then** any two complete factorizations of a are essentially the same.

30. Nov 18 – Unique factorization

Recall: if R is an integral domain, we say that **complete factorizations in R are unique, when they exist** provided for any $a \in R$ with $a \neq 0$ and $a \notin R^\times$, **if** a has a complete factorization, **then** any two complete factorizations of a are essentially the same.

Notation: Let's write "R has UCF" to abbreviate this mouthful.

Example. Recall the integral domain $R = \{a + bi\sqrt{5} : a, b \in \mathbb{Z}\}$ from Nov 13. The element 9 has two essentially different complete factorizations:

$$9 = 3 \cdot 3 \qquad \text{and} \qquad 9 = (2 + i\sqrt{5})(2 - i\sqrt{5}).$$

Hence R does not have UCF.

Recall that an element $p \in R$ of an integral domain is a **prime** if for all $a, b \in R$, $p|ab$ implies $p|a$ or $p|b$. Suppose p is prime and $p|a_1a_2 \cdots a_n = a_1(a_2 \cdots a_n)$. Then $p|a_1$ or $p|a_2 \cdots a_n$; in the latter case $p|a_2$ or $p|a_3 \cdots a_n$, etc. Hence

Lemma. *If p is a prime (in an integral domain) and $p|a_1a_2 \cdots a_n$, then $p|a_i$ for some i.*

Corollary. *Suppose R is an integral domain, $p \in R$ is a prime, and $a = q_1 \cdots q_m$ is a complete factorization of $a \in R$. Then $p|a$ iff $p \sim q_i$ for some i.*

Proof. Obviously if $p \sim q_i$ then $p|q_i$ so $p|a$. Conversely, suppose $p|a$. Then $p|q_i$ for some i, by the Lemma. Thus $q_i = pu$ for some $u \in R$. q_i is irreducible, so p or u must be a unit. p is not a unit by definition, so u is a unit. Hence $p \sim q_i$. □

Proposition. *Suppose R is an integral domain in which every irreducible element is prime. Then R has UCF.*

Proof. We repeat the proof from MATH 135/145. Suppose $a \in R$, $a \neq 0$, $a \notin R^\times$, and

$$a = p_1p_2p_3 \cdots p_n \qquad \text{and} \qquad a = q_1q_2q_3 \cdots q_m$$

where each p_i, q_j is irreducible. By assumption, each p_i is a prime. Clearly $p_1|a$, so $p_1|q_1q_2 \cdots q_m$. As p_1 is prime, the Corollary gives $p_1 \sim q_i$ for some i. We can re-order the q's so that $p_1 \sim q_1$. Then $p_1 = u_1q_1$, $u_1 \in R^\times$. Thus

$$(u_1q_1)p_2p_3 \cdots p_n = q_1q_2q_3 \cdots q_m.$$

Cancelling q_1 gives

$$u_1p_2 \cdots p_n = q_2 \cdots q_m.$$

p_2 divides the left side, so divides the right side. Hence $p_2 \sim q_j$ for some $j = 2, \ldots, m$. Again we can re-order the remaining q's and assume $p_2 \sim q_2$, say $p_2 = u_2q_2$. Then

$$u_1(u_2q_2)p_3 \cdots p_n = q_2q_3 \cdots q_m.$$

Cancelling q_2 gives

$$(u_1u_2)p_3 \cdots p_n = q_3 \cdots q_m.$$

Continuing in this way, we pair up each p_i with one of the remaining q's, until we run out of p's or q's. If we first run out of q's, i.e. $m < n$, then after m steps we will have

$$(u_1u_2 \cdots u_m)p_{m+1} \cdots p_n = 1.$$

But then $p_n|1$, which is impossible as p_n is not a unit. A similar contradiction arises if we first run out of p's. Hence $m = n$ and the two factorizations are essentially the same. □

Summary. Suppose R is an integral domain.

(1) If R does **not** have an infinite strictly increasing chain of principal ideals, then complete factorizations always exist. (Nov 17)
(2) If every irreducible in R is a prime, then complete factorizations are unique (when they exist). (Shown today)

Definition. An integral domain R is a **Unique Factorization Domain** (UFD) if (1) R does not have an infinite strictly increasing chain of principal ideals, and (2) every irreducible in R is a prime.

Example. $\mathbb{Z}$ is a UFD.

UFDs are the integral domains in which factorization works "like in $\mathbb{Z}$."

Here is one nice fact about UFDs.

Definition. Suppose R is an integral domain and $a_1, \ldots, a_n \in R$. We say that $a_1, \ldots, a_n$ are **coprime** if the only common divisors of $a_1, \ldots, a_n$ are the units in $R^\times$.

Lemma. *Suppose R is a UFD and $a_1, \ldots, a_n \in R$ with at least one $a_i \neq 0$. Then there exists $d \in R$ such that*

(1) *$d|a_i$ for each $i = 1, \ldots, n$.*
(2) *If $a_i = da_i'$ for $i = 1, \ldots, n$, then $a_1', \ldots, a_n'$ are coprime.*

[Proof on Thursday]

31. Nov 20 – UFDs and PIDs

Lemma. *Suppose R is a UFD and $a_1, \ldots, a_n \in R$ with at least one $a_i \neq 0$. Then there exists $d \in R$ such that*

(1) *$d | a_i$ for each $i = 1, \ldots, n$.*
(2) *If $a_i = da_i'$ for $i = 1, \ldots, n$, then $a_1', \ldots, a_n'$ are coprime.*

Proof sketch. The interesting case is when $a_i \neq 0$ and $a_i \notin R^\times$ for all $i = 1, \ldots, n$. Then each a_i has a complete factorization $a_i = p_{i,1}p_{i,2}\cdots p_{i,k_i}$. Let d be the "greatest common factor" of these factorizations (noting that an irreducible may may have many associates). When we divide each a_i by d to get a_i', the resulting elements $a_1', \ldots, a_n'$ have complete factorizations with no common irreducible factor. In a UFD, this means $a_1', \ldots, a_n'$ are coprime (Assignment 8). □

Recall (Nov 4) that an ideal I of a commutative ring is a **prime ideal** if $I \neq R$ and for all $a, b \in R$, if $ab \in I$ then $a \in I$ or $b \in I$. In particular, every maximal ideal is a prime ideal (see Nov 4).

Lemma. *Let R be an integral domain and $p \in R$ with $p \neq 0$. (p) is a prime ideal iff p is a prime.*

Proof. ($\Rightarrow$) Assume (p) is a prime ideal. We already know that $p \neq 0$. p cannot be a unit, since if it were, then we would have $(p) = (1)$, contradicting the assumption that $(p) \neq R$. Finally, assume $a, b \in R$ and $p | ab$. Then $ab \in (p)$. Since (p) is prime, we get $a \in (p)$ or $b \in (p)$, i.e., $p | a$ or $p | b$. Thus p is prime.
($\Leftarrow$) Proved similarly (exercise). □

Recall that a **Principal Ideal Domain** (PID) is an integral domain in which every ideal is principal.

Proposition. *Suppose R is a PID and $p \in R$ with $p \neq 0$. The following are equivalent:*

(1) *p is irreducible.*
(2) *p is a prime.*
(3) *(p) is a maximal ideal.*

Proof. (3) $\Rightarrow$ (2). If (p) is a maximal ideal, then (p) is a prime ideal, so p is prime by the previous Lemma.
(2) $\Rightarrow$ (1). Every prime is irreducible (Prop. 8.10, Nov 13).
(1) $\Rightarrow$ (3). Assume p is irreducible. Then $(p) \neq (1)$ and there is no principal ideal properly between (p) and (1). But R is a PID, so this means there is no ideal properly between (p) and (1). That means (p) is a maximal ideal. □

Here is an easy corollary that will be important in PMATH 348.

Corollary. *Suppose R is a PID and p is an irreducible element in R. Then $R/(p)$ is a field.*

Proof. (p) is maximal, so $R/(p)$ is a field (Prop. 7.12, Nov 3). □

Here is the main theorem of this section.

Theorem. *Every PID is a UFD.*

Proof. Let R be a PID. The previous proposition shows that every irreducible element of R is a prime. It remains to show that R has no infinite strictly increasing chain of principal ideals. Suppose, to the contrary, that $(a_1) \subset (a_2) \subset \cdots \subset (a_n) \subset \cdots$ is an infinite strictly increasing chain of principal ideals. Let $I = \bigcup_{n=1}^{\infty}(a_n)$. Recall that I is an ideal.

Because R is a PID, I is principal, say $I = (c)$. Then $c \in \bigcup_{n=1}^{\infty}(a_n)$, so $c \in (a_n)$ for some n. But then $(c) \subseteq (a_n)$, contradiction. □

Corollary. *If F is a field, then $F[x]$ is a UFD.*

Proof. $F[x]$ is a PID (because it has a division algorithm – Nov 11). □

Example. $\mathbb{Z}[x]$ is not a PID (Assignment 7, prob. 5). So we cannot use the above theorem to deduce that $\mathbb{Z}[x]$ is a UFD. Similarly, the ring of polynomials $F[x, y]$ in two variables is not a PID, even if F is a field (Assignment 8), so we cannot use the above theorem to prove that such polynomial rings are UFDs. Next week we will see results that imply $\mathbb{Z}[x]$ and $F[x_1, \ldots, x_n]$ are UFDs.

32. Nov 24 – Gauss' Lemma

Lemma. *Suppose R is an integral domain and $p \in R$ is a prime in R. Then p is prime in $R[x]$.*

Proof. Assume $f(x), g(x) \in R[x]$ and $p|f(x)g(x)$. Write

$$\begin{aligned} f(x) &= a_0 + a_1x + \cdots + a_mx^m \\ g(x) &= b_0 + b_1x + \cdots + b_nx^n. \end{aligned}$$

Thus

$$f(x)g(x) = c_0 + c_1x + \cdots + c_{m+n}x^{m+n} \qquad \text{where } c_k = \sum_{i+j=k} a_ib_j.$$

Since $p|f(x)g(x)$, we have $p|c_k$ for all k.

Suppose neither $f(x)$ nor $g(x)$ is divisible by p. Thus at least one coefficient of $f(x)$ and one of $g(x)$ are not divisible by p. Let r and s be the first such that $p \nmid a_r$ and $p \nmid b_s$. Let $k = r + s$ and look at c_k:

$$c_k = (a_0b_k + \cdots + a_{r-1}b_{s+1}) + a_rb_s + (a_{r+1}b_{s-1} + \cdots + a_kb_0).$$

By the choice of r and s, p divides each a_i for $i < r$ and b_j for each $j < s$. Since p also divides c_k, we get $p|a_rb_s$. As p is prime, we get $p|a_r$ or $p|b_s$, contradicting our choice of r, s. □

Lemma. *Suppose R is a UFD, $f(x), g(x) \in R[x]$, and $u \in R$, $u \neq 0$. If $u|f(x)g(x)$, then there exists a factorization $u = cd$ of u in R such that $c|f(x)$ and $d|g(x)$.*

Proof. If u is a unit (i.e., $u \in R^\times$), then we use $u = u1$. Clearly $u|f(x)$ (since $u|1$) and $1|g(x)$.

Assume u is not a unit. Because R is a UFD, u has a complete factorization

$$u = p_1p_2 \cdots p_n, \qquad \text{each } p_i \text{ irreducible.}$$

Again because R is a UFD, each p_i is prime in R and so is a prime in $R[x]$ by the previous lemma.

We have $p_1|f(x)g(x)$, so p_1 divides $f(x)$ or $g(x)$. Say $p_1|f(x)$. Let $f_1(x) \in R[x]$ be the result of dividing $f(x)$ by p_1; then

$$p_1p_2 \cdots p_n|(p_1f_1(x))g(x).$$

Cancelling p_1, we get

$$p_2 \cdots p_n|f_1(x)g(x).$$

Repeating the argument, p_2 must divide $f_1(x)$ or $g(x)$. Continuing in this way, we can "factor out" each p_i. If c is the product of the p_i's we remove from $f(x)$ and d is the product of the p_i's we remove from $g(x)$, then $cd = u$, $c|f(x)$, and $d|g(x)$. □

In the next Proposition, think of R being $\mathbb{Z}$ and F being $\mathbb{Q}$.

Proposition 9.5 (Gauss' Lemma)**.** *Suppose R is UFD and F is its field of fractions $\{n/d : n, d \in R,\ d \neq 0\}$ Let $p(x) \in R[x]$ by a polynomial of degree ≥ 1.*

Every nontrivial factorization of $p(x)$ in $F[x]$ can be essentially realized in $R[x]$, in the following sense: if $p(x) = A(x)B(x)$ is a nontrivial factorization of $p(x)$ in $F[x]$, then there exists $t \in F^\times$ such that $tA(x) \in R[x]$ and $t^{-1}B(x) \in R[x]$.

The point is that if $a(x) := tA(x)$ and $b(x) := t^{-1}B(x)$ then $p(x) = a(x)b(x)$ is a nontrivial factorization of $p(x)$ in $R[x]$.

Example. Let $R = \mathbb{Z}$, $F = \mathbb{Q}$, and $p(x) = 2x^2 + 7x + 3$. A nontrivial factorization of $p(x)$ in $\mathbb{Q}[x]$ is

$$p(x) = (x + \frac{1}{2})(2x + 6).$$

We can multiply the first factor by 2 and the second factor by $\frac{1}{2}$ to get an equivalent factorization

$$p(x) = (2x + 1)(x + 3),$$

which is a factorization in $\mathbb{Z}[x]$.

Proof of Gauss' Lemma. Each coefficient of $A(x)$ is a fraction n_i/d_i with $n_i, d_i \in R$. Let r be the product of all the denominators in $A(x)$ and let $f(x) = rA(x)$. Then $f(x) \in R[x]$ (we have "cleared the denominators"). Similarly let s be the product of the denominators in $B(x)$ and define $g(x) := sB(x) \in R[x]$. Finally let $u = rs$ and note that $u \in R$ and

$$up(x) = (rs)A(x)B(x) = f(x)g(x).$$

By the previous Lemma, there is a factorization $u = cd$ of u in R such that $c|f(x)$ and $d|g(x)$. Thus $f(x) = ca(x)$ and $g(x) = db(x)$ with $a(x), b(x) \in R[x]$. Note that $cd = rs$, so $r/c = d/s$. Let $t = r/c$. Then

$$\begin{aligned} tA(x) &= (r/c)A(x) = (1/c)rA(x) = (1/c)f(x) = a(x) \in R[x] \\ t^{-1}B(x) &= (s/d)B(x) = (1/d)sB(x) = (1/d)g(x) = b(x) \in R[x] \end{aligned}$$

as required. □

33. Nov 25 – Primitive polynomials over a UFD

Recall the statement of Gauss' Lemma from yesterday (not printed here).

The following result is particularly useful in Galois theory (PMATH 348).

Corollary. *Suppose $f(x) \in \mathbb{Z}[x]$, $\deg(f(x)) \geq 1$, and $f(x)$ is irreducible in $\mathbb{Z}[x]$. Then $f(x)$ is irreducible in $\mathbb{Q}[x]$.*

Proof. If $f(x)$ has a nontrivial factorization in $\mathbb{Q}[x]$, then $f(x)$ has a nontrivial factorization in $\mathbb{Z}[x]$ by Gauss' Lemma. □

The converse is false. For example, $6x+8$ is irreducible in $\mathbb{Q}[x]$ (every polynomial of degree 1 is irreducible), but it is reducible in $\mathbb{Z}[x]$ since $6x+8 = 2(3x+4)$ is a nontrivial factorization in $\mathbb{Z}[x]$ (neither 2 nor $3x+4$ is a unit).

Definition. Suppose R is an integral domain and $f(x) = a_0+a_1x+\cdots+a_nx^n \in R[x]$. We say that $f(x)$ is **primitive** if its coefficients $a_0, a_1, \ldots, a_n$ are coprime in R.

Corollary. *Suppose R is a UFD and F is its field of fractions. Let $f(x) \in R[x]$ with $\deg(f) \geq 1$. The following are equivalent:*

1. *$f(x)$ is irreducible in $R[x]$.*
2. *$f(x)$ is primitive in $R[x]$ and irreducible in $F[x]$.*

Proof sketch. (1) ⇒ (2) uses Gauss' Lemma to prove irreducibility in $F[x]$.

(2) ⇒ (1). Assume $f(x)$ is primitive and irreducible in $F[x]$ but is reducible in $R[x]$. Then the nontrivial factorization of $f(x)$ in $R[x]$ must be of the form $f(x) = dg(x)$ (if both factors had degrees ≥ 1 then it would be a nontrivial factorization in $F[x]$). Thus $d|f(x)$, so d divides each coefficient of $f(x)$, contradicting that $f(x)$ is primitive. □

Corollary. *Suppose R is a UFD. Every nonzero polynomial $f(x) \in R[x]$ can be factored $f(x) = dg(x)$ where $d \in R$, $g(x) \in R[x]$, and $g(x)$ is primitive.*

Example. In $\mathbb{Z}[x]$, $6x+8 = 2(3x+4)$ with $3x+4$ primitive.

Proof of Corollary. Write $f(x) = a_0 + a_1x + \cdots + a_nx^n$. By the 1st Lemma from Nov 20, there exists $d \in R$ which is a common divisor of $a_0, \ldots, a_n$ and, if $a_i = da_i'$ for $i = 0, \ldots, n$, then $a_0', \ldots, a_n'$ are coprime. Obviously $f(x) = d(a_0'+a_1'x+\cdots+a_n'x^n) = dg(x)$ and $g(x)$ is primitive. □

Crucial Lemma. *Suppose R is a UFD, $c, d \in R$ are nonzero, and $f(x), g(x) \in R[x]$ are primitive. If $(cf) \subset (dg)$, then*

1. *$(c) \subseteq (d)$,*
2. *$\deg(f) \geq \deg(g)$, and*
3. *Either $(c) \subset (d)$ or $\deg(f) > \deg(g)$.*

Remark. (cf) and (dg) are ideals in $R[x]$. (c) and (d) denote the ideals in R.

Proof. Assume $(cf) \subset (dg)$. Hence $dg(x)|cf(x)$, so

$$d|cf(x) \qquad \text{and} \qquad g(x)|cf(x).$$

The second obviously implies $\deg(f) \geq \deg(g)$, proving (2). Because $d|cf(x)$, yesterday's second lemma says that d has a factorization $d = ab$ such that $a|c$ and $b|f(x)$. Because $f(x)$ is primitive, b must be a unit. Hence (using $d = ab$) we get that d and a are associates, i.e., $d \sim a$, which implies $d|a$ (see Nov 13). As $d|a$ and $a|c$, we get $d|c$ and hence $(c) \subseteq (d)$. This proves (1).

To prove (3), assume that (3) fails, i.e., $(c) = (d)$ and $\deg(f) = \deg(g)$.

- From $(c) = (d)$, we can write $\boxed{d = cu}$ for some unit $u \in R^\times$.
- From $dg(x)|cf(x)$ and the fact that f, g have the same degree, we get $\boxed{cf(x) = e(dg(x))}$ for some $e \in R$.
- Hence $cf(x) = e(cu)g(x)$, so cancelling c we get $\boxed{f(x) = eug(x)}$.
- Hence $e|f(x)$. But $f(x)$ is primitive. Hence e is a unit.
- Hence (from the 2nd bullet) $\boxed{cf(x) \sim dg(x)}$.

But that would imply $(cf) = (dg)$, contradicting our assumption. $\square$

34. Nov 27 – The Big Theorem

Theorem 9.7. *If R is a UFD, then so is $R[x]$.*

Proof. Assuming R is a UFD, we must show that

(1) $R[x]$ has no infinite strictly increasing chain of principal ideals, and
(2) Every irreducible polynomial in $R[x]$ is prime.

(1) Assume that $(f_1) \subset (f_2) \subset \cdots \subset (f_n) \subset \cdots$ is an infinite strictly increasing sequence of principal ideals in $R[x]$.

By a Corollary from Tuesday, we can factor each $f_n(x) = c_n g_n(x)$ where $c_n \in R$ and $g_n(x)$ is primitive. Thus

$$(c_1 g_1) \subset (c_2 g_2) \subset \cdots \subset (c_n g_n) \subset \cdots$$

By the Crucial Lemma, we have

$$(c_1) \subseteq (c_2) \subseteq \cdots \subseteq (c_n) \subseteq \cdots$$

and

$$\deg(g_1) \geq \deg(g_2) \geq \cdots \geq \deg(g_n) \geq \cdots$$

and for every i,

$$(c_i) \subset (c_{i+1}) \qquad \text{or} \qquad \deg(g_i) > \deg(g_{i+1}).$$

The second option cannot happen infinitely often, since degrees are nonnegative integers. Hence beyond some point we always have the first option, meaning

$$(c_N) \subset (c_{N+1}) \subset \cdots \subset (c_{N+k}) \subset \cdots$$

But that means R has an infinite strictly increasing chain of principal ideals, contradicting that R is a UFD. This proves (1).

(2) Assume that $p(x)$ is an irreducible polynomial in $R[x]$ and $a(x), b(x) \in R[x]$ with $p(x)|a(x)b(x)$. I must show that $p(x)|a(x)$ or $p(x)|b(x)$.

By a result from Tuesday, we know that $p(x)$ is primitive in $R[x]$ and irreducible in $F[x]$, where F is the field of fractions of R.

We also know that $F[x]$ is a UFD (because F is a field, so $F[x]$ is a PID). Hence every irreducible in $F[x]$ is a prime in $F[x]$. Hence $p(x)$ is a prime in $F[x]$.

Since $a(x), b(x) \in F[x]$ and $p(x)|a(x)b(x)$, it follows that $p(x)|a(x)$ in $F[x]$ or $p(x)|b(x)$ in $F[x]$. Assume for simplicity that $p(x)|a(x)$ in $F[x]$. This means there exists $g(x) \in F[x]$ such that $a(x) = p(x)g(x)$.

Our goal is to prove $g(x) \in R[x]$, which will imply $p(x)|a(x)$ in $R[x]$. For now, however, we do not know that $g(x) \in R[x]$.

The coefficients of $g(x)$ are fractions. Let d be the product of all the denominators and let $g_1(x) = dg(x)$. Then $d \in R$ and $g_1(x) \in R[x]$ (this is "clearing denominators"). Multiplying the equation $a(x) = p(x)g(x)$ by d gives

$$da(x) = p(x)g_1(x)$$

where everything is now in $R[x]$ (or R).

Thus $d|p(x)g_1(x)$ in $R[x]$. By the second Lemma from Monday, d has a factorization $d = uv$ with $u, v \in R$, such that $u|p(x)$ and $v|g_1(x)$. But $p(x)$ is primitive, so u must be a unit, which implies $d|g_1(x)$ (in $R[x]$). Since $g_1(x) = dg(x)$, this means $g(x) \in R[x]$. □

Corollary. $\mathbb{Z}[x]$ *is a UFD.*

Proof. Because $\mathbb{Z}$ is a UFD. □

Corollary. *If R is a UFD (for example, $\mathbb{Z}$ or any field), then the ring $R[x, y]$ of polynomials over R in two variables is a UFD.*

Proof. Every polynomial in two variables, say $3x^2y + 5xy - 2xy^2 + 4x - y + 2$, can be written as a polynomial in one variable (y) whose coefficients are elements of $R[x]$. For example,

$$3x^2y + 5xy - 2xy^2 + 4x - y + 2 \;=\; (-2x)y^2 + (3x^2 + 5x - 1)y + (4x + 2).$$

Hence $R[x, y] = (R[x])[y]$. Since R is a UFD, so is $R[x]$, and hence so is $(R[x])[y]$ by two applications of the Theorem. □

Obviously we can repeat this to show that $R[x_1, \ldots, x_n]$ is a UFD for any n.

35. Dec 1 – Vandermonde determinants

Nothing in this lecture will be covered on the final exam.

Some things you didn't learn

(1) Every permutation $\sigma \in S_n$ is either **even** or **odd**, according to whether it can be written as the product of an even number of 2-cycles, or as a product of an odd number of 2-cycle.
(2) The set of even permutations in S_n is denoted A_n. It is a subgroup of S_n of index 2, called the **alternating subgroup**. $|A_n| = n!/2$.

The notion of even/odd is useful when discussing determinants. Given $\sigma \in S_n$, define

$$\operatorname{sgn}(\sigma) = \begin{cases} 0 & \text{if } \sigma \text{ is even} \\ 1 & \text{if } \sigma \text{ is odd} \end{cases}$$

Then for any matrix $A = (a_{ij})_{n\times n}$,

$$\det(A) = \sum_{\sigma\in S_n} (-1)^{\operatorname{sgn}(\sigma)} \left(\prod_{i=1}^{n} a_{i\sigma(i)} \right)$$

By last week's results, you know that the polynomial ring $\mathbb{Z}[x_1, \ldots, x_n]$ is a UFD. Today I'll show you a cute application of this result.

Definition. A **Vandermonde matrix** is any $n \times n$ matrix of the form

$$V = \begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} \\ 1 & \alpha_3 & \alpha_3^2 & \cdots & \alpha_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1} \end{bmatrix}$$

Wikipedia claims that

$$\det(V) = \prod_{1\le i<j\le n} (\alpha_j - \alpha_i).$$

Let's prove it. First, replace the α_i's with x_i's:

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

and consider the entries as elements of $\mathbb{Z}[x_1, \ldots, x_n]$. By the above formula for a determinant, we see that $\det(V)$ is an element of $\mathbb{Z}[x_1, \ldots, x_n]$.

For any $i < j$, consider the result of subtracting row i from row j:

$$V_j \;=\; \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_i & x_i^2 & \cdots & x_i^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & x_j - x_i & x_j^2 - x_i^2 & \cdots & x_j^{n-1} - x_i^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

Note that

(1) Every member of row j is divisible (in $\mathbb{Z}[x_1, \ldots, x_n]$) by $x_j - x_i$.
(2) Hence $x_j - x_i$ divides $\det(V_j)$ (in $\mathbb{Z}[x_1, \ldots, x_n]$).
(3) $\det(V_j) = \det(V)$.

Also note that the only units in $\mathbb{Z}[x_1, \ldots, x_n]$ are ± 1, so $x_j - x_i$ cannot be an associate with $x_\ell - x_k$ if $k < \ell$ but $(k, \ell) \neq (i, j)$. Also, each $x_j - x_i$ is clearly irreducible. Hence the $x_j - x_i$'s $(i < j)$ are **distinct irreducible factors** of $\det(V)$. By unique factorization, their product divides $\det(V)$, i.e.,

$$\prod_{i<j}(x_j - x_i) \mid \det(V).$$

Now let's analyze the degrees of the two polynomials. By earlier analysis, we see that every term in the sum describing $\det(V)$ has degree $0+1+2+\cdots+(n-1) = n(n-1)/2$. That is also equal to the degree of every term in $\prod_{i<j}(x_j - x_i)$ (when expanded). This implies that

$$\det(V) = c\prod_{i<j}(x_j - x_i)$$

for some $c \in \mathbb{Z}$. To evaluate c, consider the term $x_n^{n-1}x_{n-2}^{n-2}\cdots x_3^2 x_2$ (this is the product of the diagonal entries). It occurs in $\prod_{i<j}(x_j - x_i)$ with coefficient 1, and also in $\det(V)$ with coefficient 1, so $c = 1$.