MATH 235 CLASS NOTES

McGILL UNIVERSITY

NICHOLAS HAYEK

Based on lectures by Prof. Eyal Goren

CONTENTS

I Preliminaries

METHODS OF PROOF

One may show a result directly, by establishing a contradiction, proving the contrapositive, or employing induction:

1. Suppose P is false. Then, $P \implies K$, where K violates one of our assumptions. Thus, P is true.
2. To show the implication $A \implies B$, we instead show $\neg B \implies \neg A$.
3. To show that P_n is true for all $n \in \mathbb{N}$, establish that a *base case*, P_0, is true. Then, assuming that P_n holds, show that P_{n+1} also holds.
   (a) Strong induction, which assumes $P_1, ..., P_n$ to imply P_{n+1}, is logically equivalent to the following fact: for all non-empty subsets N of $\mathbb{N}$, there exists a least element $a \in N$ such that $a \leq a_i$ for all $a_i \in N$. You can also prove this principle *using* (regular) induction.

1.1 Pigeonhole Principle
Let there be n pigeonholes and greater than n pigeons. If all pigeons must be assigned to a pigeonhole, there exists at least one pigeon hole containing more than one pigeon.

Exercise: show that, for any collection C of 6 natural numbers, there exists a combination $a, b \in C$ such that $a - b$ is divisible by 5.

FUNCTION PRIMER

We define functions in the following way:

$$f : A \to B$$

where A and B are both sets. To be a function, *all* items $a_i \in A$ must be mapped to *one* item $b_i \in B$. Note the following examples.

---

1. $f : \mathbb{R} \to \mathbb{R} \quad \sqrt{x} = y$, where *both* roots are defined, is not a function.
2. $f : A \to B$ with $A := \{1, 2, 3\}, B := \{4\}$ may *only* be a function if all of A is mapped to B (not one $a \in A$ may be excluded).
3. $f : A \to \mathbb{N}$, where $\exists \ a \in A$ such that $f(a) = n_1 \in \mathbb{N}$ **and** $f(a) = n_2 \in \mathbb{N}$ is not a function.

There are three basic classes of valid functions:

1. **Injective**: if $f(x_1) = f(x_2)$, then $x_1 = x_2$.
2. **Surjective**: if, $\forall y \in Y \ \exists x \in X$ such that $f(x) = y$. In other words, all of the co-domain (output) is assigned an input.
3. **Bijective**: if f is both injective and surjective. In other words, there is a "1-to-1" correspondence between the domain and co-domain.

Exercise: If f and g are both bijective, then $h = f \circ g$ is bijective.

Measuring Infinities

Cantor defined a notion of size that can be applied to sets whose size is too large to describe with a number. This notion is called "cardinality" and is denoted by $|A|$ for the set A.

Exercise: using the Cantor-Bernstein theorem, prove that the set of points C contained within a circle of radius 1 has the same cardinality as the set of points S contained within a *square* of side length 1. What if r was arbitrary?

- A set A with a particular cardinality has the *same* cardinality as a set B *if* there exists a bijective function $f : A \to B$ (note that this is the same thing as saying that there exists a bijective function $f^{-1} : B \to A$).
- For sets A and B, $|A| \leq |B|$ if there exists an injective function $f : A \to B$.
- $|A| \geq |B|$ if there exists a surjective function between A and B.
- Cardinality preserves the transitive property of equality, that if $|A| = |B|$ and $|B| = |C|$, then $|A| = |C|$.

With respect to the above notes, when we consider an injective function, it must not also be surjective, and when we consider a surjective function, it must not also be injective (otherwise, these would just be bijective, and $|A| = |B|$.)

Proposition 1.1 Let A, B be sets. If $|A| \leq |B|$ then $|B| \geq |A|$.

Proof.

Let $A = \varnothing$. Then $|B| \geq |A|$ by definition. Now suppose that A is nonempty. Then $\exists f : A \to B$ which is injective. Define $g : B \to A$ and let $a_0 \in A$:

$$g(b) = \begin{cases} a_0 \text{ if } b \notin \text{Im}(A) \\ a \text{ if } b \in \text{Im}(A) \text{ such that } f(a) = b \end{cases}$$

Note that this is well-defined $\forall b \in B$ and surjective, so we are done. □

1.2 Cantor Bernstein

If $|A| \leq |B|$ and $|A| \geq |B|$, then $|A| = |B|$, i.e., if there exists strictly injective and surjective functions from $A \to B$, then there is a bijection between A and B.

Examples

1. Let $|\mathbb{N}| = \aleph$ (not'n: the "aleph zero"). All infinite, countable sets have carnality $\aleph$. Notably, $|\mathbb{Z}| = |\mathbb{N}| = \aleph$.
2. $|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}| = |\{(n_1, n_2) \ \forall n \in \mathbb{N}\}|$, the Cartesian product over the natural numbers.
3. $|\mathbb{N}| \neq |\mathbb{R}|$. This was shown first by Cantor using his diagonalization argument:

PROOF OF (3).

We'll assume $|\mathbb{N}| = |\mathbb{R}|$ and then work toward a contradiction.

$\implies$ $\mathbb{R}$ is countable, so let's enumerate it as follows:

$$\begin{aligned}
s_1 &= 8\ 3\ 4\ 0\ 2\ 8\ 3\ldots \\
s_2 &= 9\ 5\ 6\ 7\ 5\ 1\ 8\ldots \\
s_3 &= 2\ 9\ 5\ 6\ 2\ 1\ 2\ldots \\
s_4 &= 5\ 6\ 3\ 2\ 8\ 9\ 0\ldots \\
s_5 &= 9\ 3\ 5\ 7\ 6\ 3\ 7\ldots \\
s_6 &= 9\ 9\ 6\ 8\ 9\ 4\ 8\ldots \\
s_7 &= 6\ 5\ 3\ 7\ 4\ 9\ 8\ldots \\
&\vdots
\end{aligned}$$

Without loss of generality, we may think of any arbitrary collection of natural numbers to represent a real number in this enumeration (the decimal, too, may be anywhere or nowhere). Consider the first n real numbers in this enumeration. We then may construct an $n+1^{\text{th}}$ number which is *not* contained in the enumeration. For the above example, let's consider this hypothetical eighth number:

$$s_8 = a_1\ a_2\ a_3\ a_4\ a_5\ a_6\ a_7$$

where a_i corresponds to the i^{th} digit in the i^{th} real number:

$$\begin{aligned}
s_1 &= \mathbf{8}\ 3\ 4\ 0\ 2\ 8\ 3\ldots \\
s_2 &= 9\ \mathbf{5}\ 6\ 7\ 5\ 1\ 8\ldots \\
s_3 &= 2\ 9\ \mathbf{5}\ 6\ 2\ 1\ 2\ldots \\
s_4 &= 5\ 6\ 3\ \mathbf{2}\ 8\ 9\ 0\ldots \\
s_5 &= 9\ 3\ 5\ 7\ \mathbf{6}\ 3\ 7\ldots \\
s_6 &= 9\ 9\ 6\ 8\ 9\ \mathbf{4}\ 8\ldots \\
s_7 &= 6\ 5\ 3\ 7\ 4\ 9\ \mathbf{8}\ldots \\
\\
s_8 &= \mathbf{8}\ \mathbf{5}\ \mathbf{5}\ \mathbf{2}\ \mathbf{6}\ \mathbf{4}\ \mathbf{8}\ldots
\end{aligned}$$

Clearly, s_8 is not a member of our enumeration. Letting n be arbitrary, we now have an algorithm to generate a real number that cannot be contained in any enumeration of reals. $\lightning$

$\therefore \mathbb{R}$ is uncountable $\implies |\mathbb{R}| \neq |\mathbb{N}|$ $\square$

Cantor also devised his *Continuum hypothesis*, which states that there is no set which has a cardinality between that of the naturals and the reals. Godel and his incompleteness theorems later showed that the Continuum hypothesis could not be proven or disproved.

II Relations and Equivalency

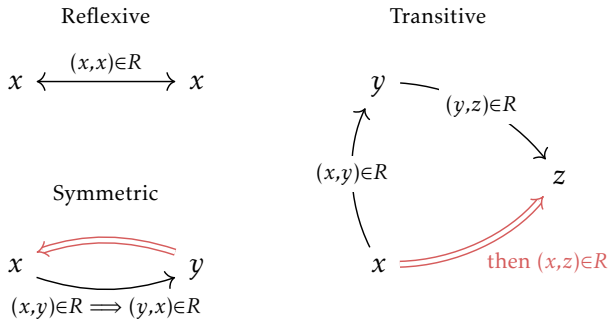
RELATIONS

Define a *relation* on a set A to be a subset $R \subseteq A \times A$. As the name suggests, relations will define various comparisons we can make between two numbers (for example, a relation defining "greater than" on $\mathbb{N}$ would contain *all* ordered pairs (a, b) where $a > b$).

At this moment, we can let R be arbitrary.

We say that an element $x \in A$ is related to y if $(x, y) \in R$, and we notate this as $x \sim y$. A relation on A is called

- **Reflexive** if $x \sim x \; \forall x \in A$
- **Symmetric** if $x \sim y \implies y \sim x \; \forall x, y \in A$
- **Transitive** if $x \sim y$ and $y \sim z \implies x \sim z \; \forall x, y, z \in A$



We also have *anti-symmetry*, where $x \sim y$ is anti-symmetric if, $\forall\; x, y$, we have

$$x \sim y \;\text{ and }\; y \sim x \implies x = y$$

Examples

- $R = A \times A$ satisfies our conditions for reflexivity, symmetry, and transitivity.
- $R = \varnothing$ is trivially symmetric and transitive (the implication of two false statements is itself true).
- $R = (a_i, a_i) \in A \times A$ is apparently reflexive, but also trivially symmetric and transitive.

A relation on the set A is called a *partial order* if it is reflexive, symmetric, and satisfies the condition

$$x \sim y \text{ and } y \sim x \implies x = y$$

Our notation can then adapt such that $x \sim y$ becomes $x \leq y$, and we then have the properties

(1) $x \leq x$ (2) $x \leq y$ and $y \leq z \implies x \leq z$ (3) $x \leq y$ and $y \leq x \implies x = y$

We call a *linear order* a partial order for which all elements x and y have either $x \leq y$ or $y \leq x$. Other names for this are *total order* and *simple order*.

Finally, an *equivalence relation* is one which is reflexive, symmetric, and transitive. This is primarily useful for identifying "like" elements (intuitively: an element is like itself, two like elements can only be so mutually, and if a member has likeness to two other members, then those two are like as well). An obvious example is the equivalence of elements in $\mathbb{R}$. Equivalence relations, and later classes, are extremely important to the study of algebra.

Examples

1. Define a *permutation* as a bijection $\sigma : A \to A$. We denote the set of all permutations of A as the set S_n, where n is the number of elements in A. The size of S_n, that is, the number of permutations of A, is always $n!$. Let's define a relation on the set of permutations such that

$$\forall \sigma, \tau \in S_n \quad \sigma \sim \tau \text{ if } \sigma(1) = \tau(1)$$

With respect to the paragraph above, that we are "identifying" members of A which are alike in their first element.

where $\sigma(1)$ is the first element of σ. Clearly, $\sigma(1) = \sigma(1)$; if σ shares τ's first element, then the converse is true; finally, if φ shares a first element with both σ and τ, then so do σ and τ. Thus, our relation is an equivalence relation.

2. Let S be the set of all sets (!). Define the following relation for all $A, B \in S$:

$$|A| = |B| \implies A \sim B$$

Once again, this is an equivalence relation.

P but not L Divisibility over $\mathbb{N}$: $a|a$ always, $a|b$ and $b|c \implies a|c$, but for $a = 5$ and $b = 3$ neither $a|b$ nor $b|a$ is true.

P and L Greater than or equal to in $\mathbb{R}$: $a \geq a$ always, $a \geq b$ and $b \geq c \implies a \geq c$, and for any $a, b \in \mathbb{R}$, either $a \geq b$ or $b \geq a$.

...with P being a partial order relation, and L a linear one.

Equivalence Classes

Let S be a set and $S_i \in S$ are subsets indexed by I. We say S is a *disjoint union* of S_i $\forall i \in I$ if $S = \bigcup_{i \in I} S_i$ and if, for $i \neq j$, we have $S_i \cap S_j = \varnothing$. We can also say that $\{S_i\}_{i \in I}$ *partitions* S.

Given an equivalence relation on A with $x \in A$, define the *equivalence class* of x, denoted $[x]$, to be the set

$$[x] = \{y \in A : x \sim y\}$$

We then have observe the following facts:

2.1 Characterization of Equivalence Classes

1. The equivalence classes of A form a partition of A
2. Any partition of A is a set of equivalence classes for a particular equivalence relation.

PROOF.

Lemma: Let X be an equivalence class and $a \in X$, then $X = [a]$

Since X is an equivalence class, $X = [x]$ for *some* chosen $x \in A$. Let $a \in X$. If $b \in [a]$ then $b \sim a$ and $a \in [x]$ also implies $a \sim x$.

$$\implies b \sim x \implies b \in [x] \implies [a] \subseteq [x]$$

As $a \sim x$, we have $x \in [a]$, so we also have $[x] \subseteq [a] \implies X = [a]$ //

We can see immediately from the lemma that every $a \in A$ is in some equivalence relation (namely, $[a]$). What we need to show, then, is that $\forall$ equivalence relations X and Y, $X \cap Y \neq \varnothing$, X and Y must be the same.

Let $a \in X \cap Y$. Then, from our lemma, $[a] = X$ and $[a] = Y \implies X = Y$, and we are done. □

With this under our belt, we can refer back to some of our examples:

> If $\sigma(1) = i$, then its equivalence class $[i]$ is all permutations such that, too, $\sigma(1) = i$. $\{[i]\}_{i \in I}$ partitions the set of all permutations. For positive integer-valued permutations, the index set I is just numbers 1,..., 9.
>
> Let an equivalence class of a set A be all sets with the same cardinality. For infinite, countable sets, the corresponding equivalence class is $[\aleph]$.

Let $\{X_i\}_{i \in I}$ be a partition of the set A. If $x \sim y$ for $x, y \in A$, then $\exists i \in I$ such that both x and $y \in X_i$. PROPOSITION 2.1

PROOF.

We can check all the conditions of equivalence relations to verify this statement:

1. Clearly, if $x \in X_i$, then $x \sim x$
2. $x \sim y$ also holds trivially, since $x \wedge y \in X_i \implies y \wedge x \in X_i$
3. Let $x \sim z$ and $y \sim z$. Then we have $x, z \in X_i$ and $y, z \in X_j$. Thus, $z \in X_i \cap X_j$. We've already shown that, if $X_i \cap X_j \neq \varnothing$, then $X_i = X_j$, which is what we have here.

   $\therefore x, y \in X_i \implies x \sim y$ □

Define the *complete set of representatives* of a set A as all $\{a_i, i \in I \subseteq A\}$ such that the set of equivalence classes is exactly $\{[a_i], i \in I\}$, with no repetitions.

For example, if S is the set of students in MATH 235, and for $x, y \in S$, $x \sim y$ if they share the same birthday, we can create an index set B of all birthdays corresponding to at least one person. Then, $\{x_b \ b \in B\}$ is the complete set of representatives of S.

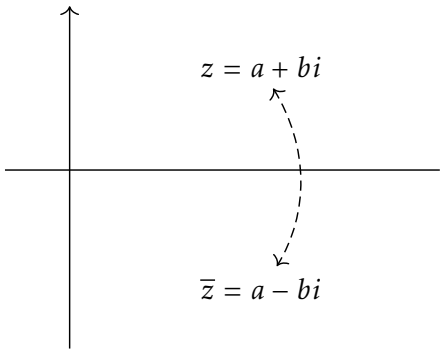
III Number Systems

COMPLEX NUMBER PRIMER

Define the set of complex numbers:

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\} \quad \text{with } i = \sqrt{-1}$$

We typically refer to a member of $\mathbb{C}$ as $z \in \mathbb{C}$. The *complex conjugate* of z, notated $\overline{z}$, has $b \to -b$. Visually, we have $\overline{z}$ reflecting z over the real axis in the complex plane.



We can also write z as a combination of its real and imaginary parts: $z = \mathfrak{Re}(z) + \mathfrak{Im}(z)i$. When adding and subtracting complex numbers, we consider their real and complex components separately, as though they were vector-valued (in many ways, they are). For example, we have:

$$[a + bi] + [c + di] = [a + c] + [b + d]i$$

Note the following identities for $z = x + yi$:

1a. $z + \overline{z} = 2\mathfrak{Re}(z)$ 1b. $z - \overline{z} = 2\mathfrak{Im}(z)i \implies x = \frac{z+\overline{z}}{2} \quad y = \frac{z-\overline{z}}{2}$

2. $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$

We define multiplication using the identity $i^2 = -1$ and the condition that the distributive property be preserved:

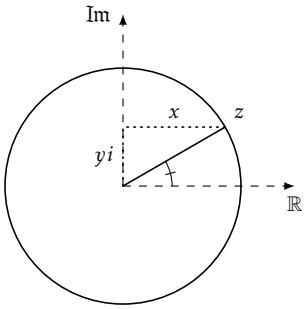
$$z_1 \cdot z_2 = (x_1 + y_1 i)(x_2 + y_2 i) = (x_1 x_2 - y_1 y_2) + (x_1 y_2 + x_2 y_1)i$$

Our basic properties of multiplication and addition are then preserved:

$$(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3) \qquad z_1(z_2 z_3) = z_1(z_2 z_3)$$
$$z_1 + z_2 = z_2 + z_1 \qquad z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$$
$$z_1 z_2 = z_2 z_1$$

for addition for multiplication

Further, we have that $\overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2}$, which is maybe unexpected. All of the above identities can be verified with some computational proofs.



Define the *magnitude* (or absolute value) of the number $z = x + yi$ to be

$$|z| = \sqrt{x^2 + y^2}$$

This implies two other identities, that

$$z \cdot \overline{z} = |z|^2 \quad \text{and} \quad |z_1 z_2| = |z_1| \cdot |z_2|$$

To round off our construction of imaginary numbers, we have the inverse: $\frac{1}{z} = \frac{\overline{z}}{|z|^2}$. One can verify that $\frac{1}{z} \cdot z = 1$

Finally, we have that complex numbers satisfy the triangle inequality, i.e.

$$|z_1 + z_2| \leq |z_1| + |z_2|$$

Polar Representations

For any point $z \in \mathbb{C}$, we can express it's coordinates in the real-imaginary plane in terms of polar coordinates, where

$$z = r\cos\theta + r\sin\theta i$$

Using trigonometric identities, we can then express the product of $z_1 z_2$ as

$$z_1 z_2 = r_1 r_2 \left[\cos(\theta_1 + \theta_2) + \sin(\theta_1 + \theta_2)i\right] \implies z_1 z_2 = (r_1 r_2, \theta_1 + \theta_2)$$

where $z_1 = (r_1, \theta_1)$ and $z_2 = (r_2, \theta_2)$. We can see from this equation that products produce a linear stretch and rotation.

Solving Polynomials in $\mathbb{C}$

We can also consider the roots of polynomials in $\mathbb{C}$:

- $x^2 + 1 = 0 \implies x = \pm i$

- For $x^2 + u = 0$, we have two solutions (unless $u = 0$). First, take $u = (-r, -\theta)$. Then $x = (\sqrt{r}, \frac{\theta}{2})$ **or** $x = (\sqrt{r}, \frac{\theta}{2} + \pi)$ may be solutions. Notice that, in the latter solution, x^2 has a rotation of $\theta + 2\pi$, where the 2π has no impact.

3.1 The Fundamental Theorem of Algebra
Any polynomial $a_n x^n + ... + a_0$ with $a_i \in \mathbb{C}$, $n > 0$, $a_n \neq 0$ has a complex root.

Example:

Let $n > 0$ with $x^n = 1$. We see from above that this equation has complex roots. In fact, for each n, there are n complex roots of one (these are called *roots of unity*).

We can express 1 in the complex plane as $(1, 2\pi k)$ in polar coordinates for any k. An n^{th} root of 1 must also be of the form $(1,$ something$)$. For this something, we have that $\frac{2\pi k}{n} \cdot n$ equals our θ expression for 1. Thus, $x = (1, \frac{2\pi k}{n})$ solves $x^n = 1$. Note that k takes values $k \in [1, n]$.

3.2 Factorizing Complex Polynomials
Let $f(x) = a_n x^n + ... + a_1 x + a_0$ be a complex polynomial of degree n. Then, we have that

$$f(x) = a_n \prod_{i=1}^{n} (x - z_i)$$

where $z_1, z_2, ..., z_n$ are the complex-valued roots of $f(x)$ such that *any* root of $f(x)$ is necessarily z_i.

De Moivre's Theorem

Let $\{z_n\}_{n \geq 1}$ be a sequence of complex numbers. We can define convergence the following way

$$\lim_{n \to \infty} |z_n - z| = 0$$

Define the complex exponential function

$$e^z = 1 + \frac{z}{1} + \frac{z^2}{2} + ... + \frac{z^n}{n} + ...$$

We can then say the following about this function:

1. This series converges absolutely for any $z \in \mathbb{C}$
2. $e^{z_1 + z_2} = e^{z_1} \cdot e^{z_2}$
3. If θ is a real number, then

3.3 Euler's Identity
$e^{i\theta} = \cos\theta + i\sin\theta$

This last fact is particularly significant. Note that $|e^{i\theta}| = 1$ always.

PROPOSITION 3.1

Let $z = e^{x+yi}$. Then we can write this number in polar coordinates as (e^x, y), using $\cos\theta$ and $i\sin\theta$ substitutions.

RINGS

Sometimes called "closed under addition/multiplication"

Consider the sets $\mathbb{Z}, \mathbb{C}, \mathbb{R}, \mathbb{N}$, etc. For any two elements in any of these sets, their product and their sum will remain in this set. For example, define $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$. Then $a + bi + c + di = (a + c) + (b + d)i \in \mathbb{Z}[i]$. These are rings.

Define a *ring* R as a set equipped with two operations:

$$\mathbb{R}\times\mathbb{R} \to \mathbb{R} : (a, b) \to a + b \qquad \text{and} \qquad \mathbb{R}\times\mathbb{R} \to \mathbb{R} : (a + b) \to a \cdot b$$

where the left operation is called "addition," and the right "multiplication." We have that the following axioms must hold.

1. Addition is commutative: $a + b = b + a \;\; \forall a, b \in R$
2. Addition is associative: $a + (b + c) = (a + b) + c$
3. There exists a zero element $\mathbb{0}$ s.t. $a + \mathbb{0} = a$
4. There exists an inverse in addition: $\forall a \in R \; \exists b \in R$ s.t. $a + b = 0$

---

5. Multiplication is associative: $a \cdot (b \cdot c) = (a \cdot b) \cdot c \; \forall a, b, c \in R$
6. There exists an identity element $\mathbb{1}$ s.t. $a \cdot \mathbb{1} = \mathbb{1} \cdot a = a$

---

7. Addition and multiplication are distributive:

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{and} \quad (b + c) \cdot a = b \cdot a + c \cdot a$$

Define $M_2(\mathbb{Z})$ and $M_2(\mathbb{R})$ as follows:

$$M_2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{with } a, b, c, d \in \mathbb{Z} \text{ or in } \mathbb{R}, \text{ respectively.}$$

Then, for both $M_2(\mathbb{Z})$ and $M_2(\mathbb{R})$, we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} a + x & b + y \\ c + z & d + w \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} ax + bz & ay + bw \\ cx + dz & cy + dw \end{pmatrix}$$

from linear algebra. We can see that these relations satisfy the axioms we established (e.g. matrix multiplication is not commutative, but addition is). Note that we do not require an inverse for multiplication.

Define a *field* to be a multiplicatively commutative, non-zero ring R such that $\forall\ x \in R$ with $x \neq 0$, $\exists\ y \in R$ such that $xy = yx = 1$. A ring may be commutative, but not a field.

The *zero ring*, $\{\mathbb{0}\}$, is not quite a field, since it is not non-zero. However, it retains all other qualities. For any element, we have $\mathbb{0} + \mathbb{0} = \mathbb{0} + \mathbb{0} = \mathbb{0} = \mathbb{0} \cdot \mathbb{0} = \mathbb{0} \cdot \mathbb{0}$, to be pedantic. Also, we have $\mathbb{1} \cdot \mathbb{0} = \mathbb{0}$ for $\mathbb{1} = \mathbb{0}$. Conversely, we have that $\mathbb{1} = \mathbb{0}$ implies that we are working with the zero ring.

Examples:

$\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}[i]$, and $\mathbb{Q}[i]$ are all commutative rings.

$\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Q}[i]$ are fields, whereas those not included from above are not.

$M_2(\mathbb{Z})$ and $M_2(\mathbb{R})$ are not commutative, and therefore not fields.

Some immediate consequences from the ring axioms:

(a) $\mathbb{0}$ is unique. In other words, if $\exists x \in R$ such that $x + a = a$, then x must be $\mathbb{0}$. Our proof is as follows:

$\mathbb{0} + x = x$ since $\mathbb{0}$ is the zero-element, and $x + \mathbb{0} = \mathbb{0}$ from above $\implies x = \mathbb{0}$

(b) $\mathbb{1}$, too, is unique, i.e. if there exists $x \in R$ such that $x \cdot a = a \cdot x = a$, then $x = \mathbb{1}$. Once again, our proof is simple:

$\mathbb{1} \cdot x = x$, since $\mathbb{1}$ is the identity element, and $1 \cdot x = \mathbb{1}$ from above $\implies x = \mathbb{1}$

(c) The element $b : a + b = \mathbb{0}$ is uniquely determined by a. We denote this b as $-a$. Note the following: $-(-a) = a$ and $-(x + y) = -x - y$.

(d) Lastly, $x \cdot \mathbb{0} = \mathbb{0}$

Subrings

Let R be a ring. We define a *subring* to be a subset $S \subseteq R$ such that:

1. $\mathbb{0}, \mathbb{1} \in S$
2. $x, y \in S \implies x + y \in S, -x \in S, x \cdot y \in S$

Check for yourself that these conditions satisfy the ring axioms. In other words, any subring is itself be a ring.

Examples:

$\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ are subrings

$\mathbb{Z} \subseteq \mathbb{Z}[i] \subseteq \mathbb{Q}[i] \subseteq \mathbb{C}$ are subrings

$M_2(\mathbb{Z}) \subseteq M_2(\mathbb{R})$ are subrings

$R = \{f : \mathbb{R} \to \mathbb{R}\}$ is a ring, where $f \in R$, $g = 0$ is the zero element, and $h = 1$ is the $\mathbb{1}$ element. Here, R is commutative, but it is not a field (prove for yourself).

$\{f : \mathbb{R} \to \mathbb{R}\}$ with f continuous is a subring.

$\{f : \mathbb{R} \to \mathbb{R}\}$ with $f(x) \in \mathbb{Z}$ is a subring.

$\{f : \mathbb{R} \to \mathbb{R}\}$ with $|f(x)| \leq 1 \ \forall x \in \mathbb{R}$ is *not* a subring.

ARITHMETIC ON INTEGERS

PROPOSITION 3.2
For example, if $a = 17$ and $b = 5$, then choose $q = 3$ so that $a - qb = 2$, which is our residue (notice that $r < |b|$ as required).

Let a, b be integers with $b \neq 0$. Then, there exists unique integers q and r such that

$$a = qb + r \quad \text{with} \quad 0 \leq r \leq |b|$$

In other words, all integers may be represented as a product of two other non-zero integers q and b, plus a residue r.

PROOF.

Assume that $b > 0$. Let $S := \{a - bx : x \in \mathbb{Z} \ \ a - bx \geq 0\}$.

Lemma: $S \neq \varnothing$. If $a \geq 0$, take $x = 0 \implies S = a$. If $a < 0$, take $x = a \implies S = a(1 - b)$, where $a < 0$, $1 - b \leq 0 \implies S \geq 0$. //

With S nonempty, take r to be the minimal element of S. Then $r = a - bq$ for some $q \in \mathbb{Z} \implies a = bq + r$. It remains to show that $0 \leq r < |b|$, which is left as an exercise. □

The proofs for these are relatively simple and are good exercise.

We say that, for any $a, b \in \mathbb{Z}$, *a divides b* if $b = ac$ for some $c \in \mathbb{Z}$. This is notated $a|b$. The following are some consequences of this definition:

(1) 0 is divisible by any integer$a$

(2) 0 only divides 0

(3) $a|b \implies a| - b$

(4) $a|b$ and $a|c \implies a|(b \pm c)$

(5) $a|b \implies a|bc$ for any $c \in \mathbb{Z}$

(6) $a|b$ and $b|a \implies a = \pm b$

PROOF OF (6).

Assume that a or $b = 0$. Then (6) holds trivially. Assume both a and b are nonzero. Then we have $a = bc$ and $b = ad$

$\implies a = acd \implies 1 = cd$, since a nonzero. Thus we have that $c = d = 1 \vee c = d = -1 \implies a = (1)b \vee a = (-1)b$ □

Let a, b be integers, with at least one nonzero. The *greatest common divisor* (or $\gcd(a, b)$, or (a, b), as you like) of a and b is the greatest positive integer that divides both a and b.

Note that, if both a and b are nonzero, then $d = \gcd(a, b) \leq \min\{|a|, |b|\}$. For a proof, take $d|a \implies a = dc$, meaning $|a| = |d||c| \geq |d| = d$. $d \leq |b|$ is shown similarly.

3.4 Bézout's Identity

Let $a, b \in \mathbb{Z}$, at least one nonzero, with $d = \gcd(a, b)$. Then:

1. $\exists u, v \in \mathbb{Z}$ such that $d = ua + vb$
2. d is the minimum positive integer of the form $ua + vb$
3. Every common divisor of a and b *itself* divides d

PROOF.

Let $S := \{ma + nb : m, n \in Z,\ ma + nb > 0\}$

Note that $S \neq \varnothing$ (one can choose $u = a$ and $v = b$, and since a^2 and b^2 are therefore positive, S contains the element $a^2 + b^2$). Choose $D \in S$ to be the minimal element of S.

Also, we have that $D|a$ and $D|b \implies D \leq d$. Let E be any common divisor of both a and b. $E|a \implies E|ua$ and $E|b \implies E|vb$, and thus $E|ua + vb$, or $E|D$. Thus, any common divisor of a and b will further divide $D \implies d|D$, and we have $D = d$. □

Example:

Consider $\gcd(7611, 592)$. As it turns out, there exists the following equation: $195 \cdot 7611 - 2507 \cdot 592 = 1$. Once cannot minimize further, so we know 1 is the greatest (and only) common divisor. But how does one come up with this?

The Euclidean Algorithm

The following will determine the gcd of any two integers a, b.

1. Let $a, b > 0$ with $a \geq b$. This assumption maintains generality, since $\gcd(a, b) = \gcd(-a, b) = \gcd(a, -b)$. If $b|a$, then clearly $\gcd(a, b) = b$, and vise-versa. Suppose not.

To demonstrate, consider gcd(48, 27):

$$\begin{aligned} 48 &= 27 * 1 + 21 \\ 27 &= 21 * 1 + 6 \\ 21 &= 6 * 3 + \boxed{3} \\ 6 &= \boxed{3} * 2 \end{aligned}$$

Our last non-zero residue was 3, and thus $(48, 27) = 3$.

2. By theorem, we have

$$\begin{aligned} &a = bq_0 + r_0 \text{ with } 0 < r_0 < b \\ &b = r_0 q_1 + r_1 \text{ with } 0 < r_1 < r_0 \\ &r_0 = r_1 q_2 + r_2 \text{ with } 0 < r_2 < r_1 \\ &\vdots \\ &r_{t-2} = r_{t-1} q_t + r_t \text{ with } 0 < r_t < r_{t-1} \\ &r_{t-1} = r_t q_{t+1} + 0 \end{aligned}$$

where r_t is the last non-zero residue. As it turns out, $r_t = \gcd(a, b)$

Proof.

Write $a = r_{-2}$ and $b = r_{-1}$. We'll prove by induction that r_t divides both r_{t-i} and r_{t-i-1} for all $0 \leq i < t + 1$ (this would imply that r_t divides both a and b).

Base case: for $i = 0$, we have $r_t|r_t$ and $r_t|r_{t-1}$. The first statement is trivial; for the second, note that $r_{t-1} = r_t q_{t+1}$.

$i \to i + 1$. From above, we have

$$r_{t-i-2} = r_{t-i-1} q_{t-i} + r_{t-i}$$

We need to show that $r_t|r_{t-i-2}$, or that $r_t|r_{t-i-1} q_{t-i} + r_{t-i}$

Since $\exists X : r_t X = r_{t-i-1} q_{t-i}$ and $\exists Y : r_t Y = r_{t-1}$, we have $r_t(X+Y) = r_{t-i-1} q_{t-i} + r_{t-1}$. Thus, r_t divides $r_{t-i-2} \implies r_t|r_{t-i}$ and $r|t_{t-i-1}$ for $0 \leq i < t+1$, or $r_t|a$ and $r_t|b$. It remains to show that, if r_t divides the "previous two" r_t's as proven, it is the greatest common divisor of a and b. □

PRIMES

Let $p \geq 2$ be a integer. We say that p is *prime* if its only positive divisors are p and 1. Numbers which are *not* prime are called composite.

Proposition 3.1 Every natural number $n \geq 2$ is a product of prime numbers (i.e. all numbers have a "prime factorization").

Proof.

We'll show by induction. Base case: for $n = 2$, n is clearly a product of primes

$n \to n+1$: Let n be prime. If we take $n+1$ to be prime, then we are done. If $n+1$ is not a prime, then $n+1 = rs$ with $1 < s \leq n$. Since r and s are less than $n+1$, we know that *they* are products of primes.

$\therefore$ $n+1$ is a product of primes.

Thus, if n is positive, then $n = \varepsilon p_1 p_2 ... p_k$ with $\varepsilon = 1$, and if n is negative, then $n = \varepsilon p_1 p_2 ... p_k$ with $\varepsilon = -1$. □

Sieve of Eratosthenes Detour

Let $n \geq 2$ be an integer. If n is not prime, then n is divisible by some prime $1 < p \leq \sqrt{n}$. As proof, consider that a non-prime integer may be written as a product of 2 or more primes. If all of these primes are greater than $\sqrt{n}$, then their product is *greater* than n. Thus, we require at least one prime that is $\leq \sqrt{n}$.

From here, one can algorithmically "cancel out" all non-primes from a set of numbers. The following theorem will strengthen the proposition above.

3.4 Fundamental Theorem of Arithmetic
Let n be a non-zero integer. Then the prime factorization

$$n = \varepsilon p_1 p_2 ... p_k$$

is unique.

PROOF OF THEOREM PROPOSITION 3.2

Let $p \geq 2$ be an integer. Then the following are equivalent:

(a) p is prime (b) If $p|ab$, where a, b are non-zero integers, then $p|a$ or $p|b$.

PROOF OF PROPOSITION.

Assume (2). Suppose $p = st$, a product of two integers. WLOG, assume s, t are positive (else, let $s = -s$, $t = -t$). Then by (2) we have $s = \omega p$. Then $p = st = p\omega t$, which are all positive integers.

$\implies \omega = t = 1$, and then $p = s \implies p$ has no non-trivial factors, and is prime.

Assume (1). Given that $p|ab$, we need to show that either $p|a$ or $p|b$. If $p|a$, then we are done. If $p \nmid a$, we need to show $p|b$.

Since $p \nmid a$, we have that $\gcd(p, a) = 1$ (note that p is prime to see why). Then we have $1 = up + va$ for some $u, v \in \mathbb{Z}$. Multiplying by b, we get $b = upb + vab$.

$p|ab \implies p|vab$ and $p|p \implies p|upb \implies p|(vab + upb) \implies p|b$ □

We'll show the main theorem by induction. Consider the following two prime factorizations: $n = \varepsilon p_1 p_2 ... p_k$ and $n = \mu q_1 q_2 ... p_l$

We will need to show that $p_i = q_i$, $\mu = \varepsilon$, and $k = l$. Let $i = 1$ be our base case: we have $n = \varepsilon p_1 = \mu q_1$, implying n is itself a prime number, with a fixed sign, and so $p_1 = q_1$ trivially.

For $q_1 \leq p_1$, the proof will be identical.

$i - 1 \to i$: we have that $k \geq 1$ and $l \geq 1$. Assume WLOG that $p_1 \leq q_1$

$$p_1|n \implies p_1|q_1 q_2 ... q_l$$

and thus p_1 divides *some* q_i. However, since the chosen q_i must be prime, we can conclude that $p_1 = q_i$

Further, we have $p_1 \leq q_1 \leq q_i$ with $p_1 = q_i \implies p_1 \leq q_1 \leq p_1 \implies p_1 = q_1$.

$\implies \frac{n}{p_1} = p_2 ... p_k = q_2 ... q_l$. However, we have that $p_2 = q_2, ..., p_k = q_l$ by our induction hypothesis. □

Proposition 3.4 We will consider some immediate corollaries, the first of which being that there are infinitely many prime numbers. One proves this with a diagonalization argument.

Proposition 3.5 For non-zero integers a and b, $a|b$ IFF

$$a = p_1^{a_1} p_2^{a_2} ... p_n^{a_n} \implies b = p_1^{a_1'} p_2^{a_2'} ... p_n^{a_n'} \quad \text{with } a_i' \geq a_i$$

where p_i are distinct primes, and a_i *may* be zero (this allows us to disregard primes which make up b, but not a).

Proof.

($\implies$) Suppose we can write a and b in this fashion. Then we have $b = p_1^{a_1} p_2^{a_2} ... p_n^{a_n} p_1^{a_1' - a_1} p_2^{a_2' - a_2} ... p_n^{a_n' - a_n} = a \cdot p_1^{a_1' - a_1} p_2^{a_2' - a_2} ... p_n^{a_n' - a_n}$, so $a|b$.

($\impliedby$) Let $a|b$. Since $a|b$, $b = ak$ for an integer k.

We order primes as follows: let $p_1, ..., p_l$ be primes contained in the unique factorization of k that are also contained in that of a. Label their exponents a_i and b_i for a and k, respectively. Let $p_{l+1}, p_{l+2}, ..., p_m$ be primes in the unique factorization of a that are not in that of k. Finally, let $p_{m+1}, ..., p_n$ be primes in the factorization of k which are not in that of a, and label their exponents b_i as well. We can then write:

$$a = p_1^{a_1} ... p_l^{a_l} p_{l+1}^{a_{l+1}} ... p_m^{a_m} p_{m+1}^0 ... p_n^0 \quad \text{and} \quad k = p_1^{b_1} ... p_l^{b_l} p_{m+1}^{b_{m+1}} ... p_n^{b_n}$$

Thus, $b = ak = p_1^{a_1+b_1}...p_l^{a_l+b_l}p_{l+1}^{a+l}...p_m^{a_m}p_{m+1}^{b_{m+1}}...p_n^{b_n}$. This is precisely $p_1^{a_1'}...p_n^{a_n'}$. □

PROPOSITION 3.6

Let $a = p_1^{a_n}...p_n^{a_n}$ and $b = p_1^{b_1}...p_n^{b_n}$ for distinct primes p_i (their exponents may be 0). Then

$$\gcd(a,b) = p_1^{\min(a_1,b_1)}...p_n^{\min(a_n,b_n)} \quad \text{and} \quad \operatorname{lcm}(a,b) = p_1^{\max(a_1,b_1)}...p_n^{\max(a_n,b_n)}$$

3.5 Fundamental Theorem of Algebra for $\mathbb{Q}$
For any non-zero rational number $q \in \mathbb{Q}$, we have the unique prime factorization $q = p_1^{a_1}p_2^{a_2}...p_m^{a_m}$

EXAMPLE

We'll show that $\sqrt{2}$ is irrational using the fundamental theorem. Let $\sqrt{2}$ be rational, with $q = \frac{a}{b}$ for non-zero integers a, b $\in \mathbb{Z}$

Then $\sqrt{2} = p_1^{a_1}p_2^{a_2}...p_m^{a_m}$, and this is unique.

$\implies 2 = p_1^{2a_1}p_2^{2a_2}...p_m^{2a_m}$, and this too is a unique factorization.

However, since 2 is prime, we know that its factorization is $2 = 2^1$

$2 = p_1^{2a_1} \implies p_1 = 2 \quad \text{and} \quad 2a_1 = 1$

$\implies a_1 = 1/2$, which is a contradiction. ↯

CONGRUENCES

Fix $n > 1$, an integer. Define the relation on $\mathbb{Z}$

$$x \sim y \text{ if } n|(x-y)$$

We also notate $x \equiv y \pmod{n}$ or $x \equiv y$, and say that x is *congruent* to y. For example, if $n = 2$, then $x \equiv y$ if they have the same parity. Note that $x \equiv y$ is an equivalence relation:

Reflexive: $x - x = 0$, and $n|0$, so $x \equiv x$

Symmetric: $n|(x-y) \implies n|-(x-y)$, so $n|y-x$

Transitive: Given that $x \equiv y$ and $y \equiv z$, we have $n|(x-y)+(y-z)$, and thus $n|x-z$

We can also describe this relation's set of representatives, i.e. the set of elements which disjointly and completely defines all equivalence classes. It is precisely the set $\{0, 1, ..., n-1\}$.

As proof, note that all integers x can be written as $qn + r$, with $0 \leq r < n \implies x - r = qn \implies n|x - r$, or simply $x \equiv r$. We can thus deduce that, for any $x \in \mathbb{Z}$, it will be congruent to $r \in \{0, 1, ..., n-1\}$. To show their respective classes must be disjoint, assume that $x \equiv r_1$ and $x \equiv r_2$, with $0 \leq r_1 < r_2 < n$ as required. Then $n|x - r_1$ and $n|x - r_2 \implies n|r_2 - r_1$. But $r_2 - r_1 \neq 0$, and $n > r_2 - r_1$, so n cannot divide this $\lightning$

Instead of denoting these equivalence classes $[0], [1], ..., [n-1]$, write $\overline{0}, \overline{1}, ..., \overline{n-1}$. Define the set of all equivalence classes for $x \equiv y (\text{mod } n)$ as $\mathbb{Z}/n\mathbb{Z}$. Define further, under this set, the following for addition and multiplication:

$$\overline{i} + \overline{j} = \overline{i+j} \quad \text{and} \quad \overline{i}\,\overline{j} = \overline{ij}$$

In this case, $\mathbb{0} = \overline{0}$, $\mathbb{1} = \overline{1}$, and $-\overline{i} = \overline{-i}$ (the inverse of addition). Let's play around with $\overline{i}$ for a moment:

Examples

With $n = 13$, we'll compute $\overline{5} \cdot \overline{6} - \overline{5}$. We first have that $\overline{5} \cdot \overline{6} = \overline{30}$. For some $x \in \mathbb{Z}$, if we have $13|x - 30$, then $13q + 30 = x$ for an integer q, but this also holds for $13(q+2) + 4 = x$.

This way of writing preserves our condition that $r < n = 13$, and thus $\overline{30} = \overline{4}$. In any case, if $r \geq n$ or $r < 0$, $\overline{r}$ can be reduced to the remainder of r when divided by n.

We then have only to compute $\overline{4} - \overline{5}$, or $\overline{4} + \overline{-5} = \overline{-1} = \overline{12}$, which is our final answer.

Importantly, $\mathbb{Z}/n\mathbb{Z}$ is a commutative ring. A proof for this begins by showing that our stated definitions are well-defined, i.e. that $x \equiv x'$ and $y \equiv y'$ should imply that $\overline{x+y} = \overline{x'+y'}$ and similarly $\overline{xy} = \overline{x'y'}$. Then, of course, one has to trudge through demonstrating the 7 axioms.

Here are a few more examples of modular arithmetic (what we were doing in the example above–some call it "clock arithmetic"), with $n = 4$:

The inclusion of 4 in the right-most columns is pedantic; to be clear, $\overline{4} = \overline{0}$.

+	0	1	2	3	4
0	0	1	2	3	0
1	1	2	3	0	1
2	2	3	0	1	2
3	3	0	1	2	3
4	0	1	2	3	0

and

×	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	0
2	0	2	0	2	0
3	0	3	2	1	0
4	0	0	0	0	0

Define a *zero divisor* as an element $x \in R$, with $x \neq 0$, such that $\exists\ y \neq 0$ with either $xy = 0$ or $yx = 0$.

Suppose we have that R is a commutative ring. If R contains any zero divisors, then R is not a field. As proof, assume that this is not true, and pick $x \neq 0$ and $y \neq 0$ with $xy = 0$. Since R is a field, we have that $\exists x^{-1}$ with $xx^{-1} = 1$. Thus, $(xx^{-1})y = y$ but $x^{-1}(xy) = 0$ ↯

PROPOSITION 3.7

Using this fact, we can prove that $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime.

PROPOSITION 3.8

($\implies$) Suppose that $n > 1$ is not prime. Then we have that $n = ab$ for integers $a, b \in [2, n-1]$. Thus, $\overline{a} \neq 0$ and $\overline{b} \neq 0$, but $\overline{a} \cdot \overline{b} = \overline{ab} = \overline{n} = \overline{0}$. Thus, R has zero divisors, and it cannot be a field. For the case $n = 1$, note that $\mathbb{Z}/1\mathbb{Z}$ will have 1 element, i.e. $\mathbb{Z}/1\mathbb{Z}$ is the zero ring, and is not a field.

(PARTIAL) PROOF.

We denote the ring $\mathbb{Z}/n\mathbb{Z}$ as $\mathbb{F}_p$ when $n = p$ is prime. This ring has exactly p elements.

FERMAT'S LITTLE THEOREM

Let p be a prime number, and let $a \not\equiv 0 (\text{mod } p)$. Then we have that

$$a^{p-1} \equiv 1 (\text{mod } p)$$

To calculate 2^{100} (mod 13), we have that $2^{100} = 2^{96} \cdot 2^4 = (2^{12})^8 \cdot 2^4$. From Fermat's, this is simply $1^8 \cdot 2^4 = 2^4$ (mod 13), which is just 3.

EXAMPLE

IV Polynomial Arithmetic

RINGS OF POLYNOMIALS

Define the *ring of polynomials* over a particular ring R as

$$R[x] = \{a_n x^n + ... + a_1 x + a_0 : a_i \in R\}$$

with $n \geq 0$. Note that, since n may be 0, we can construct the *zero polynomial*, where both n and $a_0 = 0$. Addition is well defined on this ring, with

$$(a_n x^n + ... + a_1 x + a_0) + (b_m x^m + ... + b_1 x + b_0)$$
$$= a_n x^n + ... + (a_m + b_m)x^m + ... + (a_1 + b_1)x + (a_0 + b_0)$$

assuming $n \leq m$ WLOG. Similarly, multiplication is given by

$$(a_n x^n + ... + a_1 x + a_0)(b_m x^m + ... + b_1 x + b_0)$$
$$= c_{n+m} x^{n+m} + ... + c_1 x + c_0 \quad \text{with} \quad c_i = a_0 b_i + a_1 b_{i-1} + ... + a_i b_0$$

We say that a polynomial $f(x)$ is *monic* if $a_n = 1$. $f(x)$ is said to be of *degree* n if $a_n = 0$. And finally, define the *constant polynomial* when f has degree 0, i.e. $f(x) = a$ with $a \neq 0$.

Though it won't be shown, $R[x]$ with addition and multiplication as defined above is commutative ring. Further, $\mathbb{0}$ is the zero polynomial, $\mathbb{1}$ is the constant polynomial 1, and $-f$ is simply $-(a_n x_n + ... + a_1 x + a_0)$.

DIVISION OF POLYNOMIALS

PROPOSITION 4.1 Let $R[x]$ be such that the multiplication of any two non-zero elements is non-zero (this is called an *integral domain*, and defines most rings we've considered before). Then if $f(x)$ and $g(x) \in R[x]$ are non-zero, then we have that

$$\deg[f(x)g(x)] = \deg[f(x)] + \deg[g(x)]$$

PROOF. Suppose $\deg(f) = n$ and $\deg(g) = m$. We have that $f(x) = a_n x^n + ... + a_0$ and $g(x) = b_m x^m + ... + b_0$. Then we have that $f(x)g(x) = a_n b_m x^{m+n} + \text{LOT}$. Since R is an integral domain, $a_n b_m \neq 0$, so $\deg(fg) = m + n = \deg(f) + \deg(g)$, and we are done. □

"lower-order terms"

4.1 Division of Polynomials
Let $f(x)$, $g(x)$ be two polynomials in $\mathbb{F}[x]$, with $g(x) \neq 0$. Then there always exist two *unique* polynomials $q(x)$ and $r(x)$ in $\mathbb{F}[x]$ such that

$$f(x) = q(x)g(x) + r(x)$$

with $r(x)$ either 0 or with $\deg(r) < \deg(g)$

As we've done before, define $f|g$ for functions in $\mathbb{F}[x]$ as the property $f(x) = q(x)g(x)$ for some non-zero $q(x) \in \mathbb{F}[x]$. The following still hold: PROPOSITION 4.2

1. $f|g \implies f|-g$
2. $f|g \implies f|gh$ for any $h \in \mathbb{F}[x]$
3. $f|g$ and $f|h \implies f|g \pm h$

We define the GCD of two not-both-zero polynomials f and g as the *monic* polynomial h of largest degree for which $h|f$ and $h|g$. This is unique.

4.2 Bézout's for Polynomials
Let f, g be not-both-zero polynomials with a greatest common divisor h. Then there exists functions $u, v \in \mathbb{F}$ such that

$$h = uf + vg$$

Further, h is the smallest-degree monic polynomial satisfying this equation.

As with integers, if d is the GCD of f and g, then any common divisor h divides d. COROLLARY 4.2.1

PROOF OF COROLLARY.
Since $h|f$ and $h|g$, write $hq_1 = f$ and $hq_2 = g$ for polynomials q_1, q_2. We can write the GCD of f and g as $d = uf + vg$ for some polynomials u, v. Then $d = uhq_1 + vhq_2 = h(uq_1 + vq_2)$. Thus, $h|d$. □

Euclidean Algorithm for Polynomials

Just as we've had before for integers, there is a Euclidean algorithm for finding the GCD of two functions f and $g \in \mathbb{F}[x]$. Let $g(x) = a_n x^n + \text{LOT}$.

If $g|f$, then $a_n^{-1} g(x)$ is the GCD of f and g. Suppose not, and define indefinitely:

We multiply by a_n^{-1} since the GCD of polynomials is defined to be monic. A similar constant is factored in after the Euclidean algorithm.

$$\begin{aligned} f(x) &= q_0(x)g(x) + r_0(x) \\ g(x) &= q_1(x)r_0(x) + r_1(x) \\ r_0(x) &= q_2(x)r_1(x) + r_2(x) \\ &\vdots \\ r_{t-2}(x) &= q_t(x)r_{t-1}(x) + \boxed{r_t(x)} \\ r_{t-1}(x) &= q_{t+1}(x)\boxed{r_t(x)} \end{aligned}$$

PROPOSITION 4.3 Let $r_t(x) = c_m x^m + ... + c_m$.The GCD of f and g is $c_m^{-1} r_t(x)$.

ASSOCIATES

$\mathbb{F}^\times$ denotes the real members of $\mathbb{F}$, used to differentiate polynomials in $\mathbb{F}[x]$ with their coefficients, for example.

Let $\mathbb{F}$ be a field, $f, g \in \mathbb{F}[x]$ be two non-zero polynomials. We say that f and g are *associates* if $\exists \alpha \in \mathbb{F}^\times$ such that $\alpha f = g$. This is an equivalence relation on the set of polynomials. As proof, see that $\mathbb{F}$ being a field $\implies \alpha^{-1}$ exists.

A similar thought: if p is a prime number, for $m|p$, either $m = \pm 1$ or $m = \pm p$.

A non-constant polynomial f, with $\deg(f) > 0$, is called *irreducible* if, for any g with $g|f$, g is either an associate of $\mathbb{1}$ or of f Suppose that $\deg(f) \geq 1$. Then the following are equivalent:

PROPOSITION 4.4 (1) f is irreducible (2) $f|gh \implies f|g$ or $f|h$.

This is precisely an analog for what we did with prime numbers, and the proof, too, is similar:

PARTIAL PROOF ($\implies$) Suppose f is irreducible and $f|gh$. If $f \nmid g$, then $\gcd(f, g) = 1$. We can then write $1 = uf + vg$ for some $u, v \in \mathbb{F}[x]$. Thus, $h = ufh + vgh$, but $f|ufh$, and $f|ghu$, so $f|h$. □

The following is a lemma to an upcoming theorem. Let f be any non-zero polynomial in $\mathbb{F}[x]$. Then f can be written as

$$f = cf_1 f_2 ... f_n \quad \text{with all } f_i \in \mathbb{F}[x] \text{ irreducible and monic, with } c \in \mathbb{F}^\times$$

PROOF. By induction, suppose $\deg(f) = 0$. Then f is a constant, and $f = f$ ✓

$\deg(n) \to \deg(n+1)$: case (1): If f is irreducible, $\exists c$ such that $f = cf_1$ with f_1 monic and irreducible.

Case(2) If f is reducible, write $f = f_1 f_2$ with $\deg(f_1) < \deg(f)$, $\deg(f_2) < \deg(f)$. Then each f_1, f_2 can be written as $c_1 p_1 ... p_a$ and $c_2 p_{a+1} ... p_b$, and thus $f_1 f_2 = c_1 c_2 p_1 ... p_b$, and we are done. □

4.3 Unique Factorization for Polynomials
Let $f \in \mathbb{F}[x]$ be a non-zero polynomial. Then we have that

$$f = cp_1^{a_1} \dots p_r^{a_r}$$

for $c \in \mathbb{F}^\times$ and p_i monic, distinct, and irreducible polynomials $\forall\ i$. Moreover, c and all p_i are *uniquely determined*.

Some important results follow. Suppose f and $g \in \mathbb{F}[x]$ are nonzero. Then $f|g$ if and only if COROLLARY 4.3.1

$$f(x) = cf_1(x)^{a_1'} \dots f_r(x)^{a_r'} \quad g(x) = df_1(x)^{a_1} \dots f_r(x)^{a_r}$$

where $c, d \in \mathbb{F}^\times$ and all f_i are irreducible monic. Lastly, we have that $0 \leq a_i' \leq a_i$

($\implies$) We can easily find and h such that $g = fh$ using the form above: $h = dc^{-1} f_1(x)^{a_1 - a_1'} + f_r(x)^{a_r - a_r'}$. Thus, we conclude $f|g$ PROOF.

($\impliedby$) Assume now that $f|g$. We then have $g = fh$. We can write the following WLOG:

$$f = cf_1^{a_1'} \dots f_s^{a_s'} \qquad \text{and} \qquad h = ef_1^{b_1} \dots f_s^{b_s} f_{s+1}^{a_{s+1}} \dots f_r^{a_r}$$

Then we have that $g = cef_1^{a_1' + b_1} \dots f_s^{a_s' + b_s} f_{s+1}^{a_{s+1}} \dots f_r^{a_r}$. One lets $d = ce$, $a_i = a_i' + b_i$, and we are done. □

If f, g are non-zero polynomials with $f = cf_1^{a_1} \dots f_r^{a_r}$ and $g = df_1^{b_1} \dots f_r^{b_r}$, $c, d \in \mathbb{F}^\times$, and $a_i \in \mathbb{Z}$, then the GCD of f and g can be written as follows: COROLLARY 4.3.2

$$\gcd(f, g) = f_1^{\min\{a_1, b_1\}} \dots f_r^{\min\{a_r, b_r\}} \quad \text{and} \quad \operatorname{lcm}(f, g) = f_1^{\max\{a_1, b_1\}} \dots f_r^{\max\{a_r, b_r\}}$$

where the LCM is the smallest monic polynomial c such that $f|c$ and $g|c$.

Let $f = (x+1)(x+2)$ and $g = x(x+1)^2$. Consider the minimum and maximum exponents of these common irreducible polynomials (note: all linear polynomials are irreducible). Then $\gcd(f, g) = (x+1)$ and $\operatorname{lcm}(f, g) = x(x+1)^2(x+2)$. EXAMPLE

IDENTIFYING IRREDUCIBLE POLYNOMIALS

Just as we asked for integers and primes: how do we tell if a polynomial is irreducible? While there is no sure algorithm to answer this question, the following come in handy:

1. If $f \in \mathbb{F}[x]$ has a root, then it is reducible, and write $f(x) = (x - \alpha)g(x)$. The converse for this is not necessarily true: if we have a reducible polynomial, is *still* may have no roots in $\mathbb{F}$. If f is irreducible, then f has no roots in $\mathbb{F}$.

2. Any linear polynomial, $ax + b$, $a \neq 0$, is irreducible.

★ 3. If $f(x) \in \mathbb{F}[x]$ has degree 2 or 3, then f is reducible IFF f has root in $\mathbb{F}$.

4. In $\mathbb{C}$, the only irreducible polynomials are the linear polynomials. We saw previously that for any $f \in \mathbb{C}[x]$

$$f(x) = c \prod_{i=1}^{\deg(f)} (x - \alpha_i)$$

and the result follows.

5. In $\mathbb{R}$, any irreducible polynomial has either degree 1 or 2.

6. The number of roots of $f \in \mathbb{F}[x]$ is at most $\deg(f)$.

Proof of ★. ($\implies$) If $f(\alpha) = 0$, then $f(x) = (x - \alpha)g(x)$ for some $g(x) \in \mathbb{F}[x]$, and is thus reducible. ($\impliedby$) If f is reducible, we have $f(x) = g(x)h(x)$. Then $\deg(f) = \deg(g) + \deg(h)$. Since $\deg(f) = 2 \vee 3$, assume WLOG that $\deg(h) = 1$. Then $h(x) = ax + b$ for some $a, b \in \mathbb{F}$. Then $\alpha = -ba^{-1}$ is a root of h and thus f. □

Identifying roots of $f(x)$

Let $f(x) \in \mathbb{F}[x]$. The criterion outlined in (1) and (3) from above will be particularity useful. We'll focus a bit on how to pinpoint whether a polynomial has a root or not, which will often indicate whether it is reducible or not.

4.4 Characterization of Roots in $\mathbb{Q}[x]$
Let $f(x) = a_n x^n + ... + a_1 x + a_0$ be a non-constant polynomial with integer coefficients. If f has a root, $q = \frac{s}{t}$, where q is a reduced fraction, then

$$t | a_n \quad \text{and} \quad s | a_0$$

Proof. We have $f\left(\frac{s}{t}\right) = \left(\frac{s}{t}\right)^n a_n + \left(\frac{s}{t}\right)^{n-1} a_{n-1} + ... + \left(\frac{s}{t}\right) a_1 + a_0 = 0$. We can multiply by b^n to yield

$$\underbrace{s^n a_n + s^{n-1} t a_{n-1} + ... + s t^{n-1} a_1}_{\text{Divisible by } s} + a_0 t^n = 0 \quad s^n a_n + \underbrace{s^{n-1} t a_{n-1} + ... + s t^{n-1} a_1 + a_0 t^n}_{\text{Divisible by } t} = 0$$

Thus, $s | a_0 t^n \implies s | a_0$ and $t | s^n a_n \implies t | a_n$ □

Consider $f = x^{p-1} - 1$. Fermat's Little Theorem ensures that $a^{p-1} \equiv 1$ for any $a \not\equiv 0 \in \mathbb{Z}/p\mathbb{Z}$, so the elements $\overline{a}$ for $a \in [1, p-1]$ are roots of the polynomial $x^{p-1} - 1$.

Since there are $p-1$ elements in this set, there are at least $p-1$ roots of f. Further, since x^{p-1} has degree $p-1$, we conclude each root $\overline{a}$ has multiplicity 1, and these are the only roots.

$\implies$ multiplying by x, all elements of $\mathbb{Z}/p\mathbb{Z}$ are roots of $x^p - x$: PROPOSITION 4.4

$$x^p - x = \prod_{a=0}^{p-1} (x - \overline{a})$$

4.5 Existence of roots in $\mathbb{Z}/p\mathbb{Z}$ Suppose $f(x) \in \mathbb{Z}/p\mathbb{Z}[x]$ is a non-zero polynomial. Then, f has a root in $\mathbb{Z}/p\mathbb{Z}$ IFF $\gcd(f, x^p - x) \neq 1$. Thus, if $\gcd(f, x^{p-1} - 1) \neq 1$, f has a root.

PROOF.

($\implies$) Suppose $f(a) = 0$ for some $a \in \mathbb{Z}/p\mathbb{Z}$. Then $(x-a)|f(x)$ and $(x-a)|x^p - x$, using Prop. 4.4. Thus, $\gcd(f, x^p - x)$ has degree at least $(x - a)$, and thus $\neq 1$.

($\impliedby$) Now suppose that $h(x) = \gcd(f, x^p - x) \neq 1$. Then $h(x)|x^p - x = \prod_{a=0}^{p-1} (x - \overline{a})$. Since h has a unique factorization, we conclude that $h = \prod_{a \in I} (x - \overline{a})$, for some subset $I \in [0, p-1]$.

Since $h(x)r(x) = f(x)$, when $x = a \in I$, $h(a) = 0$ and thus $f(a) = 0$. □

If $f \in \mathbb{R}[x]$ is a polynomial of odd degree, then f has a root in $\mathbb{R}$. PROPOSITION 4.5

PROOF.

Let $f = a_n x^n + ... + a_0$, and choose $a_n > 0$ WLOG. For some large N we can guarantee that $f > 0$, and likewise for some large, negative $\tilde{N}$, $f < 0$. It follows by intermidate value theorem that $f = 0$ for some $x \in [\tilde{N}, N]$. □

V Rings

IDEALS

Recall that a ring R is a non-empty set equipped with operations addition and multiplication:

$$(x, y) \to x + y \quad \text{and} \quad (x, y) \to xy$$

In the future, assume that any ring R is commutative, i.e. $xy = yx$.

An *ideal* of ring R is a subset $I \subseteq R$ such that

(1) $\mathbb{0} \in I$ (2) $a, b \in I \implies a + b \in I$ (3) $a \in I, r \in R \implies ra \in I$ and $ar \in I$

One notates $I \triangleleft R$ for "I is an ideal of R." Typically, $\mathbb{1} \notin I$, since we have that $\mathbb{1} \in I \iff I = R$. One concludes that I is not typically a subring.

The sets $\{\mathbb{0}\}$ and R are called the "trivial ideals" of R. Note that, if R is a division ring (i.e. $\exists a^{-1} : aa^{-1} = \mathbb{1}$), then any non-zero ideal is $I = R$. As proof, see that if $a \neq 0 \in I$, then $aa^{-1} \in I$, so $\mathbb{1} \in I$. The result follows.

Let R be a commutative ring. Then we define $(r) := \{ra : a \in R\} = \{ar : a \in R\}$, where $r \in R$, to be the *principal ideal ring* of R. This set is sometimes denoted by rR, Rr or $\langle r \rangle$. One can verify that (r) is an ideal:

1. Since $\mathbb{0} \in R$ and $ra \in R$ for any $a \in R$, $r\mathbb{0} = \mathbb{0} \in (r)$
2. $ra_1 + ra_2 = r(a_1 + a_2)$. Since R is closed by addition, $r(a_1 + a_2) \in (r)$.
3. For any $ra \in (r)$, $s \in R$, $r(as) \in (r)$

5.1 Ideals of $\mathbb{Z}$
Every ideal of $\mathbb{Z}$ is a principal ideal (m), notated $m\mathbb{Z}$, for some integer $m \in \mathbb{Z}$. The complete list of ideals is precisely $(0), (1), (2), (3), \ldots$

Proof. Suppose that I is a non-zero ideal of $\mathbb{Z}$. Then we can always find a positive element $a \in I$ (otherwise, pick a negative a'; then $a'(-1) \in I$ is positive). Choose the minimal element $i \in I$. Clearly $(i) \subseteq I$ (one can check the axioms). Let $j \notin I$. Then $j = iq + r$, where $r < i$. We have then that $r = j - iq$ is an element of I *less than* i, which is a contradiction. Let $r = 0$. Then $j = iq$, and since $q \in \mathbb{Z}$, $j \in (i)$. The minimal positive element of (i) is i for all $i \in \mathbb{Z}$, thus $\{(i)\ \forall i \geq 0\}$ are a unique collection of all the ideals of $\mathbb{Z}$. □

5.2 Ideals of $\mathbb{F}[x]$
All ideals of $\mathbb{F}[x]$ can be written as (f) for a unique polynomial $f \in \mathbb{F}[x]$. Furthermore, $f \sim g \iff (f) = (g)$.

PROOF. Consider an ideal $I \subseteq \mathbb{F}[x]$. Choose the polynomial $f \in I$ with minimal degree. Recall that an associate of f is a polynomial αf for any $\alpha \in \mathbb{F}^\times$. Clearly $\alpha f \in I$, and since $\deg(f) = \deg(\alpha f)$, one could just as easily choose αf as the minimal degree element. From here, conclude that $f \sim g \implies (f) = (g)$. Furthermore, if $(f) = (g)$, then f and g's minimal degree polynomials are such that $f = g$, $f = g\alpha$, or $g = f\alpha$ for any α. Thus $(f) = (g) \implies f \sim g$.

Consider now (f). Clearly (f) is a subset of I. Suppose $g \in I$ but $g \notin (f)$. We can write $g = fq + r \implies r = g - fq$. Thus $r \in I$, but $\deg(r) < \deg(f)$, which is a contradiction. Set $r = 0$. Then $g = fq$, and thus $g \in (f)$. We conclude that $(f) = I$ for the ideal I whose minimal degree polynomial is f. □

For any two elements $r, s \in R$, we say $r \sim s$, or r and s are *associates*, if $(r) = (s)$. For two polynomials, our definition of associativity remains unchanged.

HOMOMORPHISMS

PROPOSITION 5.1 Let R and S be commutative rings. The function $f : R \to S$ is called a *ring homomorphism* if (1), (2), and (3) hold $\forall x, y \in R$. (i), (ii), (iii) follow from these axioms:

1. $f(\mathbb{1}_R) = \mathbb{1}_S$	i. $f(\mathbb{0}_R) = \mathbb{0}_S$
2. $f(x + y) = f(x) + f(y)$	ii. $-f(x) = f(-x)$
3. $f(xy) = f(x)f(y)$	iii. $f(x - y) = f(x) - f(y)$

PROOF. (i): We have that $f(\mathbb{0}_S) = f(\mathbb{0}_S + \mathbb{0}_R) = f(\mathbb{0}_S) + f(\mathbb{0}_R)$. On adds $-f(\mathbb{0}_S)$ to both sides to yield $\mathbb{0}_S = f(\mathbb{0}_R)$

(ii): From above, $\mathbb{0}_S = f(\mathbb{0}_R) = f(x + (-x)) = f(x) + f(-x)$. Adding $-f(x)$ to both sides yields $f(-x) = -f(x)$

(iii): The proof here follows from (ii). □

PROPOSITION 5.2 Define the image of the homomorphism f to be $\mathrm{Im}(f) = \{f(r) : r \in R\}$. This is a subring of S.

PROOF.
1. We have $\mathbb{0}_R, \mathbb{1}_R \in R$. Thus $f(0_R) = \mathbb{0}_S$ and $f(\mathbb{1}_R) = \mathbb{1}_S \in \text{Im}(f)$
2. Let $x_1, x_2 \in \text{Im}(f)$. Then $\exists r_1, r_2 \in R : f(r_1) = x_1$ and $f(r_2) = r_2$. Note that $r_1 + r_2 \in R$. We have $f(r_1 + r_2) = f(r_1) + f(r_2) = x_1 + x_2$, so $x_1 + x_2 \in \text{Im}(f)$
3. Similarly, $f(r_1 r_2) = f(r_1)f(r_2) = x_1 x_2$, so $x_1 x_2 \in \text{Im}(f)$.
4. $\exists r \in R : f(r) = x$. Then $f(-r) = -f(r) = -x$, so $-x \in \text{Im}(f)$ □

PROPOSITION 5.3
PROOF.

One checks individually that

(1) $\mathbb{0}_S, \mathbb{1}_S \in \text{Im}(f)$
(2) $x_1, x_2 \in \text{Im}(f) \implies x_1 + x_2 \in \text{Im}(f)$
(3) $x_1, x_2 \in \text{Im}(f) \implies x_1 x_2 \in \text{Im}(f)$
(4) $x \in \text{Im}(f) \implies -x \in \text{Im}(f)$

Let $f : R \to S$ be a homomorphism. Define the *kernel* of f to be

$$\ker(f) = \{r \in R : f(r) = \mathbb{0}_S\} = f^{-1}(\mathbb{0}_S)$$

The kernel of f is an ideal of R. Moreover, f is injective $\iff \ker(f) = \{\mathbb{0}_R\}$, and $f(x) = f(y) \iff x - y \in \ker(f)$

To show that $\ker(f)$ is an ideal of R, we first need to show $\mathbb{0}_R \in \ker(f)$. See that $f(\mathbb{0}_R) = \mathbb{0}_S$, so this is true. Now to show $a, b \in \ker(f) \implies a + b \in \ker(f)$: we have that $f(a + b) = f(a) + f(b) = \mathbb{0}_S + \mathbb{0}_S = \mathbb{0}_S$. Finally, to show $ra \in \ker(f)$ for any $r \in R$: $f(ar) = f(a)f(r) = \mathbb{0}_S f(r) = \mathbb{0}_S$, and we are done.

f is injective iff $f(x) = f(x') \iff x = x'$. Thus $f(x) - f(x') = \mathbb{0}_S \iff x - x' = \mathbb{0}_R$. Let $r = x - x'$. $f(r) = \mathbb{0}_S \iff r = \mathbb{0}_R$, and we are done. For the last claim, see that $f(x) = f(y) \iff f(x) - f(y) = \mathbb{0}_S \iff f(x - y) = \mathbb{0}_S \iff x - y \in \ker(f)$. □

Example:

Let $n \geq 1$ be some integer and define $f : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, $f(a) = \overline{a}$ (i.e. maps a to its congruence/residue class). We have that f is a homomorphism. Checking all conditions:

1. $f(1) = \overline{1}$, $1 = \mathbb{1}_{\mathbb{Z}}$, $\overline{1} = \mathbb{1}_{\mathbb{Z}/n\mathbb{Z}}$ ✓
2. $f(x + y) = \overline{x + y} = \overline{x} + \overline{y} = f(x) + f(y)$ ✓
3. $f(xy) = \overline{xy} = \overline{x}\,\overline{y} = f(x)f(y)$ ✓

Observe that $\ker(f) = \{r : f(r) = \overline{r} \equiv \mathbb{0}_{\mathbb{Z}/n\mathbb{Z}} = 0 \bmod n\} = \{r : r \equiv 0 \bmod n\}$. This is just the set of integer multiples of n, i.e. $nr\ \forall r \in \mathbb{Z}$, sometimes denoted $n\mathbb{Z}$

COSETS

All our work on $\mathbb{Z}/n\mathbb{Z}$ has been, more or less, a particular case of the general theory of cosets. Let R be a commutative ring and I be an arbitrary ideal of R. Define the following relation on R:

$$x \sim y \iff x - y \in I$$

For example, when $R := \mathbb{Z}$ and $I := n\mathbb{Z}$, $x \sim y$ if $x - y \in n\mathbb{Z}$, or if $x - y$ are multiples of n, or $n|x - y$. This is exactly $x \equiv y \bmod n$.

$x \sim y$ has the following properties: PROPOSITION 5.4

1. $x \sim y$ is an equivalence relation
2. Every equivalence class is of the form $x + I := \{x + t : t \in I\}$ for a particular $x \in R$
3. $x + I = y + I \iff x - y \in I$
4. Either $(x + I) \cap (y + I) = \varnothing$ or $x + I = y + I$

PROOF. For (1): see that $x - x = \mathbb{0} \in I$, so $x \sim x$. Then assume $x \sim y \implies x - y \in I$. Since any element in I multiplied by an element in R is in I, $\mathbb{1}(x - y) = y - x \in I \implies y \sim x$. Finally, take $x \sim y$ and $y \sim z$. Then $x - y \in I$ and $y - z \in I$, and we conclude $x - y + y - z = x - z \in I \implies x \sim z$.

For (2): consider $x + I = \{x + t : t \in I\}$ for $x \in R$. If $y \in x + I$, we have that $y = x + t$ for some t. Then $x - y = x - (x + t) = -t \in I \implies x \sim y$. Now assume that $x \sim y$. Then $x - y := s \in I$. We say $y = x + (y - x) = x + s \implies y \in x + I$. We conclude that $x + I$, $x \in R$, describes all equivalence classes.

For (3): $x + I$ is the equivalence class of x, and $y + I$ is the equivalence class of y. Since $x + I = y + I$, $x \sim y$, so $x - y \in I$.

For (4): Follows from the fact that equivalence classes form a partition of R.

Define the ring R/I, "R mod I," to be the set of equivalence classes as defined above. This ring is commutative, and addition and multiplication are defined as follows:

$$(x + I) + (y + I) := (x + y) + I$$

$$(x + I)(y + I) := xy + I$$

We will notate $\overline{x} = x + I$. This may feel familiar: $\overline{x} + \overline{y} = \overline{x + y}$ and $\overline{xy} = \overline{xy}$. Furthermore, $\mathbb{0} = \overline{0}$ and $\mathbb{1} = \overline{1}$ for this ring.

5.3 Mapping $R \to R/I$
Let R be a commutative ring, R/I defined as above. The function $\pi : R \to R/I$ $\pi(x) = \overline{x}$ is a surjective homomorphism. Furthermore, $\ker(\pi) = I$

PROOF. To check that π is a homomorphism, see that $\pi(1) = \overline{1}$; $f(a + b) = \overline{a + b} = \overline{a} + \overline{b} = f(a) + f(b)$; finally, $f(ab) = \overline{ab} = \overline{a}\overline{b} = f(a)f(b)$. This is surjective too: choose $\overline{r} = r + I$. By definition, $r \in R$.

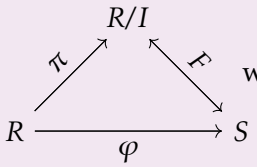
Now, $\ker(\pi) = \{r \in R : \pi(r) = \overline{0}\} = \{r \in R : r + I = I\} = \{r \in R : r \in I\}$ □

Corollary 5.3.1 Any ideal is the kernel of some ring homomorphism. As proof, we see that, for any $I \triangleright R$, $\pi : R \to R/I$ has $\ker(\pi) = I$.

Isomorphisms

For two rings R and S, one says that R is *isomorphic* to S if there exists a bijective ring homomorphism between the two. This is notated $R \cong S$.

5.4 First Isomorphism Theorem
Let $\varphi : R \to S$ be a surjective ring homomorphism. Let $I = \ker(\varphi)$. Then R/I is isomorphic to S.



where $\pi : \alpha \to \overline{\alpha}$ and F is a bijective homomorphism

5.5 Chinese Remainder Theorem
Let m, n be integers such that $\gcd(m, n) = 1$. Then

$$\mathbb{Z}/mn\mathbb{Z} \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$$

Proof. We'll use the first isomorphism theorem. Let $\varphi : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ with $\varphi(a) = (a \bmod m, a \bmod n)$. One checks that this is a surjective homomorphism. Also, $\ker(\varphi) = \{a : (a \bmod m, a \bmod n) = (0, 0)\} = \{a : a = mk_1, a = nk_2\} = \{a : n|a, m|a\}$. Since m, n are relatively prime, this is $\{mnk : k \in \mathbb{Z}\} = mn\mathbb{Z}$.

By FIT, $\mathbb{Z}/\ker(\varphi) \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \implies \mathbb{Z}/mn\mathbb{Z} \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ □

Quotient Rings

Let $\mathbb{F}$ be a field, $\mathbb{F}[x]$ be the ring of polynomials in $\mathbb{F}$, and $f(x) \in \mathbb{F}[x]$ some non-constant, irreducible polynomial. We define the ring $\mathbb{F}[x]/\langle f(x)\rangle$ to be set of cosets generated by the principal ideal $\langle f(x)\rangle$.

5.6 Elements of $\mathbb{F}[x]/\langle f(x)\rangle$
The cosets of the ring $F[x]/\langle f(x)\rangle$ can be all be written as $\overline{g(x)}$ for some unique polynomial with $0 \leq \deg(g) < \deg(f)$

PROOF. Let $h(x) \in \mathbb{F}[x]$ with $h(x) = q(x)f(x) + r(x)$. The degree of r is less than f. Furthermore, $h(x) - r(x) = f(x)q(x)$, and $f(x)q(x) \in \langle f(x)\rangle$, so $\overline{h} \equiv \overline{r}$. Thus, any arbitrary coset will be congruent to $\overline{r}$ for some $r : \deg(r) < \deg(f)$. □

5.7 Size of $\mathbb{F}[x]/\langle f(x)\rangle$
Let $f(x)$ be non-constant and irreducible in $\mathbb{F}[x]$. Then $\mathbb{F}[x]/\langle f(x)\rangle$ is a field with q^n elements, where q is the number of elements in $\mathbb{F}[x]$ and $n = \deg(f)$.

PROOF. We know that $\mathbb{F}[x]/\langle f(x)\rangle$ is a commutative ring, and, since $\overline{1} \notin \langle f(x)\rangle$, $\overline{1} \neq \overline{0}$. Consider a non-zero element $\overline{g} : \overline{g} \neq \overline{0}$. We have that $f \nmid g$, since, otherwise, $g \in \langle f\rangle \implies \overline{g} = \overline{0}$. Since f is irreducible, we conclude that $\gcd(f, g) = 1$, so $\exists u(x), v(x) : 1 = uf + vg$.

$\implies \overline{1} = \overline{u(x)f(x)} + \overline{v(x)g(x)}$, since $u(x)f(x) \in \langle f\rangle$, so $\overline{uf} = \overline{0}$. Thus, $\overline{1} = \overline{v(x)}\,\overline{g(x)}$. We then have a multiplicative inverse for g, so $\mathbb{F}[x]/\langle f(x)\rangle$ is a field.

We know that $\mathbb{F}[x]/\langle f(x)\rangle \subseteq \{b_{n-1}x^{n-1} + ... + b_0 : b_i \in \mathbb{F}\}$, i.e. the set of polynomials g with $\deg(g) < \deg(f) = n$. To show that all polynomials of this form are unique cosets, let g, h have degrees less than f, and suppose $\overline{g} = \overline{h}$. Then, $g - h \in \langle f\rangle \implies \deg(g - h) \geq n$, but this can't happen, since $\deg(g) < n$ and $\deg(h) < n$ (unless, of course, $g = h$). Thus, all polynomials of the form $b_{n-1}x^{n-1} + ... + b_0$ are unique cosets. If there are q elements of $\mathbb{F}$, this leaves us with q^n possible polynomials. □

Examples:

1. Consider $f \in \mathbb{F}_2$ and the ring $\mathbb{F}_2[x]/\langle x^2 + x + 1\rangle$. It's members are simply the remaining polynomials of the form $ax + b$ in $\mathbb{F}_2$, i.e. $\{0, 1, x, x + 1\}$. The following describe addition and multiplication in the quotient ring:

+	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

×	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	$x+1$	1
$x+1$	0	$x+1$	1	x

2. Suppose we want to find a finite field with 25 elements. This is 5^2, so $\mathbb{F}/f$ will work if $\deg(f) = 2$ is monic-irreducible and $\mathbb{F}$ has 5 elements, i.e. $\mathbb{F}_5/\langle x^2 - 2\rangle$.

5.8 Roots in Larger Fields Let $g(x) \in \mathbb{F}[x]$ be a non-constant polynomial. Then there is a field $L \supseteq \mathbb{F}$ such that $g(x)$ always has a root in L.

Proof.

One assumes WLOG that g is irreducible. Otherwise, $g'|g$ for some irreducible polynomial, and a proof of the theorem for *irreducible* polynomials $\implies g'(l) = 0 \implies f(l) = 0$.

Let $L := \mathbb{F}[x]/\langle g(x)\rangle$. This is a field. Consider the natural map $\varphi : \mathbb{F} \to \mathbb{F}[x]/\langle g(x)\rangle$ with $\varphi(\alpha) = \overline{\alpha}$. One can check that this is an injective ring homomorphism, so the image of $\varphi = \{\overline{\alpha} : \alpha \in \mathbb{F}\} = \mathbb{F}$ is a subring of $L \implies L \supseteq \mathbb{F}$.

Let $g(x) = a_n x^n + ... + a_0 : a_i \in \mathbb{F}$. Note that $a_i = \overline{a_i}$ in L. Then when $x = \overline{x}$, we have $g(\overline{x}) = \overline{g(x)}$, i.e. the ideal generating L, so $\overline{g(x)} = \overline{0}$. Thus, $g(x)$ has the root $\overline{x}$ in $\mathbb{F}[x]/\langle g(x)\rangle$. □

VI Groups

Reminiscent of rings, a group G is any non-empty set equipped with a closed operation $G \times G \to G : (a, b) \to ab$.

Group Axioms:	**Consequences:**
Associativity: $a(bc) = (ab)c$	$ab = ac \implies b = c$
Neutral Element: $\exists \mathbb{1}_G : a\mathbb{1}_G = \mathbb{1}_G a = \mathbb{1}_G$	$\mathbb{1}_G$ is unique
Inverse Element: $\forall a \in G \, \exists b \in G : ab = ba = \mathbb{1}_G$	The inverse of a is unique

Though inverses are well defined, we never notate them $\frac{1}{a}$, but instead write a^{-1}. Another consequence of the axioms is $(a_1 a_2)^{-1} = a_2^{-1} a_1^{-1}$.

Define for $a \in G$, $n \in \mathbb{Z}$:

$$a^n := \begin{cases} \mathbb{1}_G & \text{if n=0} \\ \underbrace{a \cdot a \cdot \ldots \cdot a}_{n \text{ times}} & \text{if } n > 0 \\ \underbrace{a^{-1} \cdot a^{-1} \cdot \ldots \cdot a^{-1}}_{n \text{ times}} & \text{if } n < 0 \end{cases}$$

Our axioms imply that $a^{n+m} = a^n a^m$ for any $a, m \in \mathbb{Z}$. Some things clearly do *not* hold for groups. For example, we do not have a $\mathbb{0}$ element, nor an addition operation to accompany multiplication. Usually, we also have that groups are not commutative, but they may be, and in that case we call them *abelian groups*.

When a group is abelian, and rings with addition always are, we sometimes write $\times$ as $+$. This group is still equipped with only one operation, though.

Examples:

1. The trivial group G has one element, i.e. $G = \{\mathbb{1}\}$
2. $G = \mathbb{Z}$ or $\mathbb{Z}/n\mathbb{Z}$, where addition is the operator, is an abelian group.
3. If $\mathbb{F}$ is a field, it is especially a ring, and so $(\mathbb{F}, +)$ is an abelian group, but $(\mathbb{F}, \times)$ is too, since multiplication is commutative.
4. Consider the 2 by 2 matrix ring denoted by $M_2(R) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with matrix addition and multiplication as usual. If $\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) \neq 0$, then inverses exist, so $M_2(R)^\times := \{M \in M_2(R) : ad - bc \neq 0\}$ is a non-abelian group under multiplication.

5. If $\mathbb{F}$ is a field, $M_2(\mathbb{F})^\times$ is a group as above. However, it is still non-abelian. This group is sometimes denoted $GL_2(\mathbb{F})$.

First properties and types of groups

One sometimes denotes this $H < G$.

Define a *subgroup* $H \subseteq G$ with the following properties:

1. $\mathbb{1} \in H$ 2. $a, b \in H \implies ab \in H$ 3. $a \in H \implies a^{-1} \in H$

In kind, "using addition" is only a notational change.

Define a *cyclic subgroup* $\langle g \rangle := \{g^n : n \in \mathbb{Z}\}$ with $g \in G$. In the event that our group is equipped with addition, then this definition changes to $\langle g \rangle := \{gn : n \in \mathbb{Z}\}$. Proving this is indeed a subgroup of G is straightforward using the closure property of multiplication. One calls a group *cyclic* if it is its own cyclic subgroup, i.e. G can be written as $\langle g \rangle$ for some $g \in G$ (this element is called the *generator*).

As an example, see that $\mathbb{Z}$ and $\mathbb{Z}/n\mathbb{Z}$ are cyclic under addition. An important observation is that all cyclic groups are abelian, and thus all non-abelian groups are non-cyclic.

Define the *order* of G, denoted $|G|$, and $\#G$ less often, to be the number of elements in G. We also have a notion of order for any $g \in G$, which is the minimal $n \in \mathbb{N}$ such that $g^n = 1$. One notates $\text{ord}(g) = n$. If no such n exists, write $\text{ord}(g) = \infty$.

Examples:

1. The order of $\mathbb{Z}$ is clearly infinite. The order of any positive element $k \in Z$ is *also* infinite. As proof, see that $nk = 0 \implies n = 0$, remembering that we are working with addition.

2. Let μ_n, for some fixed $n \in \mathbb{N}$, be the set of n^{th} roots of unity. $(\mu_n, \times)$ is a cyclic group with n elements, and $\mu_n = \langle e^{\frac{2\pi i}{n}} \rangle$

3. For $M_2(\mathbb{F}_2)$ as defined above, has six elements, and thus has order 6. All of its elements have a finite order, and, for instance, $\text{ord}\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right) = 2$.

Proposition 6.1 Our first big proposition about groups is that $\text{ord}(g) = |\langle g \rangle|$.

Proof. If $|\langle g \rangle|$ is finite, since one writes g^r for *all* integers, and especially positive integers, there are elements in this set such that $g^a = g^b$, with $a > b > 0$. Then $g^{a-b} = g^{b-b} = g^b g^{-b} = 1$. Thus we've found a positive $a - b$ with $g^{a-b} = 1$, so $\text{ord}(g)$ must be finite. By contrapositive, $\text{ord}(g) = \infty \implies |\langle g \rangle| = \infty$.

Now fix $\text{ord}(g) = n$. Let $a \in \mathbb{Z}$. By residue division, one writes $g^a = g^{qn+r} = (g^n)^q g^r = g^r$. We conclude that all g^a are equal to g^r for some $r < n$. Clearly, one can cook up an a such that its residue mod n is r for any r they'd like, so $|\langle g \rangle|$ has at most n elements. To show this is exactly the case, see that $g^a \neq g^b$ for $a, b < n$. Otherwise, $g^{a-b} = 1$, and $\text{ord}(g) < n$, which is a contradiction. □

Permutations and Cycles

Define a permutation $\sigma : [1, n] \to [1, n]$ which reorders the elements in $\{1, 2, ..., n\}$ to $\{\sigma(1), \sigma(2), ..., \sigma(n)\}$. There are $n!$ ways to permute the set $[1, n]$ (or any set with n elements), and all permutations functions must be bijective.

We call the set S_n the symmetry group, and it contains the set of permutations of $[1, n]$, written $S_n := \{\sigma : [1, n] \to [1, n], \sigma \text{ bijective}\}$. With its operation being compositions (of permutations), S_n is a non-abelian group for $n \geq 3$, and abelian otherwise.

S_2 contains the permutation which swaps $\{a, b\}$ and the identity, which commute. S_1 follows similarly.

Checking the group axioms, for permutations σ, ρ, we have that $\sigma \circ \rho$ is bijective $\implies \in S_n$, and S_n is closed under multiplication. Then, see that compositions are associative. The neutral permutation τ that sends elements $a_i \to a_i$ satisfies $\sigma \circ \tau = \tau \circ \sigma = \sigma$. Finally, since σ is bijective, so is σ^{-1}, so this is in S_n. As stated above, $|S_n| = n!$

When one writes out a particular permutation, the most explicit form is a table, where the top entry in a column is mapped to the entry below it:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 2 & 1 & 5 \end{pmatrix} \qquad \sigma(\text{Id}) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 2 & 1 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}$$

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

This notation has obvious limitations and redundancies, so we distill these charts into *cycles*, which store information about the closed "chains" of movement in the permutation. They are written $(a\ b\ c\ ...y\ z)$ for the integers in $[1, n]$, and can be understood as: $a \mapsto b, b \mapsto c, ..., y \mapsto z, z \mapsto a$. Sometimes a permutation requires that multiple cycles be strung together, so in truth $\{a...z\}$ are subsets of $[1, n]$. Finally, see that the inverse of $\sigma = (a\ b...z)$ is just the cycle in reverse, $(z\ y...b\ a)$.

Proving this is good exercise.

Examples:

In the examples given above, $\sigma = (4\ 1)(2\ 3)$. When an element is fixed, (5) in this case, it can be omitted from the cycle. The second line expresses $(3\ 2)(1\ 3)(2\ 3\ 1) = \text{Id}$. When multiplying a chain of cycles, do so in pairs, right to left.

We'll consider (1 5 2 4)(2 6 4)(3 4 7). The last two terms are (2 6 4)(3 4 7). Let's start at 3 in the right cycle. This will map to 4, which then is hijacked by the left cycle and redirected to 2, which now goes to 6. Thus we have (3 2 6...). Continuing on the left, we see that 6 goes to 4, but the right cycle maps this to 7, so 6 maps to 4, which is mapped to 7. Finally, we get (3 2 6 4 7). See that 7 goes to 3 in the right cycle, which is consistent with this answer.

The "method," using the logic just used, is to start in the right cycle until you count an element that appears on the left. Redirect this element to the element the left cycle maps it to, and continue on the left. Now, when you count an element that appears on the right, count it, and continue on the right. Repeat this until you have a closed loop.

We can now find (1 5 2 4)(3 2 6 4 7) without justifying each step: (3 4 7)(6 1 5 2).

For this one, note that if two cycles are disjoint, they are commutative.

For additional practice, find for yourself that (7 2)(9 1)(1 5 9)(2 8) = (1 5)(2 8 7)

Proposition 6.2

If we have a permutation in S_n, this can be written always as a product of disjoint cycles. Furthermore, we can characterize the order of $\sigma \in S_n$, i.e. the number of compositions one needs for $\sigma \circ \sigma \circ ... \circ \sigma = \text{Id}$, as $\text{lcm}(a_1, ..., a_n)$, where a_i is the length of each disjoint cycle τ_i (this is also the order of the cycle).

Partial Proof.

No proof will be given for the first claim. Let σ be a permutation of disjoint cycles $\tau_1, ..., \tau_n$. We write σ^k is $\underbrace{\tau_1...\tau_n\tau_1...\tau_n...\tau_1...\tau_n}_{k \text{ times}}$, but since all τ_i, τ_j are commutative by their disjointedness, and τ_i is commutative with itself, one can rearrange: $\sigma^k = \underbrace{\tau_1...\tau_1}_{k \text{ times}} \underbrace{\tau_2...\tau_2}_{k \text{ times}} ..... \underbrace{\tau_n...\tau_n}_{k \text{ times}}$. Thus, we have that $\sigma^k = \tau_1^k...\tau_n^k$.

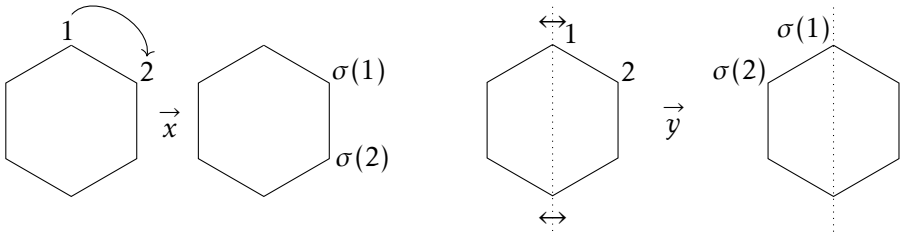
Lemma: If $g \in G$ with $\text{ord}(g) = a$, then $g^k = 1$ iff $a|k$.
As proof, write $k = qa + r$ by residue. Then $g^k = (g^a)^q g^r = g^r$. Then $g^r = 1$ with $r < a$, which is contradictory unless $r = 0 \implies k = qa \implies a|k$

Then $\sigma^k = 1 \iff \tau_1^k = ... = \tau_n^k = 1 \iff a_1|k, ..., a_n|k \iff \text{lcm}(a_1, ..., a_n)|k$. With $\text{ord}(\sigma)$ being the *minimal* k satisfying this, we conclude $\text{ord}(\sigma) = \text{lcm}(a_1, ..., a_n)$. □

Symmetries and Dihedral Groups

Define the *Dihedral group* D_n to be the group of symmetries of a regular n-gon (which is defined for $n \geq 3$), and let $x, y \in D_n$ be two of its elements, where x is a rotation about the center of the n-gon by $360/n$ degrees, and y is the reflection across its line of symmetry (one can always choose the y-axis, with the n-gon oriented appropriately).

We have that the order of x is n (composing n $360/n$-degree rotations will land you where you started), and the order of y is 2 (flip twice).



PROPOSITION 6.3

Additionally, every symmetry $\sigma \in D_n$ can be completely characterized by where it takes two adjacent lattices, here labeled 1 and 2. Notationally, we have $\sigma = \tau \iff \sigma(1) = \tau(1)$ and $\sigma(2) = \tau(2)$. We can identify all the elements of D_n, too:

$$D_n = \{x, x^2, ..., x^{n-1}, y, xy, x^2y, ..., x^{n-1}y, \mathrm{Id}\}$$

with D_n non-abelian with $2n$ elements.

PROOF.

We know that if $\sigma \in D_n$, then $\sigma(1)$ and $\sigma(2)$ are next to each-other. Thus, if $\sigma(1) = 1 + a$, then $\sigma(2) = 1 + a$ or a. $\implies \sigma = x^a$ or $x^a y$.

For D_n, we have the property* that $xy = yx^{-1}$, which comes from rearranging $xyxy = 1$. By induction, one can prove $x^a y = yx^{-a}$ in generality (good practice). Then, to show that D_n is non-abelian, take $xy = yx \implies yx^{-1} = yx \implies x^2 = 1$, and this is clearly not true in generality. □

Visualize what these mean in terms of rotations and reflections, and note that $0 = n$. A chart of symmetries looks like:

	1	2
x^a	$1 + a$	$2 + a$
$x^a y$	$1 + a$	a

As practice, what is $x^3yxyx^2yx^4$ in D_5?

This evaluates to y, with ample use of $\star$.

COSETS FOR GROUPS

PROPOSITION 6.4

Let $H < G$. Define a *left coset* of G in G to be a subset $gH \subseteq G$ with $gH := \{gh : h \in H\}$. These cosets will form an equivalence class for the relation

$$x \sim y \text{ if } y^{-1}x \in H \text{ with } x, y \in G$$

Thus, 2 cosets are either disjoint or equal under this relation. We also write

$$xH = yH \iff y^{-1}x \in H \iff x^{-1}y \in H \iff \exists h \in H : x = yh$$

and $\quad xH = H \iff x \in H$

PROOF.

First, to show $x \sim y$ is an equivalence relation, we have (1) $x^{-1}x = \mathbb{1} \in H$, so $x \sim x$; (2) if $x \sim y$, then $y^{-1}x \in H$. Since H is closed under the inverse, $(y^{-1}x)^{-1} = x^{-1}y \in H \implies y \sim x$; (3) if $x \sim y$ and $y \sim z$, write $y \sim x$ and $z \sim y$, so $x^{-1}yy^{-1}z = x^{-1}z \in H \implies z \sim x \implies x \sim z$.

To show that the left cosets form equivalence classes, suppose $x \sim y$. Then $y \sim x$, and $x^{-1}y \in H \implies xx^{-1}y \in xH \implies y \in xH$. Similarly, if $y \in xH$, then $y = xh$ for some $h \in H$, and then write $x^{-1}y = h \implies x^{-1}y \in H$, so $x \sim y$. We conclude that xH are equivalence classes under $x \sim y$, so they are either disjoint or equivalent.

Our first line of IFF statements follows from the fact that xH and yH are equivalence classes, except for the last statement, which requires a touch of manipulation: $y{-}1x \in H \iff \exists h \in H : y^{-1}x = h \iff \exists h \in H : x = yh$.

Finally, $xH = H$ follows from the line above with $y := 1$. □

6.1 Lagrange's Theorem

Let G be a finite group and $H < G$ be a subgroup. Define the *index* of H with respect to G, notated $[G : H]$, to be the number of distinct left cosets of G w.r.t H. We then have

$$[G : H]|H| = |G| \quad \text{and} \quad |H| \text{ divides } |G|$$

Corollary 6.1.1

An immediate corollary is that $\text{ord}(g) \mid |G|$, where $\text{ord}(g) = |\langle g \rangle|$ as usual, since $\langle g \rangle$ is a subgroup of G for $g \in G$.

Proof.

We need to show that all cosets have the same size, and this size is $|H|$. Then, since left cosets are equivalence classes, they form a partition of G, and thus $[G : H]|H| = |G|$. Also, $\#H|\#G$ follows immediately.

Notation: $|S| = \#S$ denotes the size of a set (# is usually reserved for finite sets).

Define $f : H \to xH$ with $f(h) = xh$ for some fixed $x \in G$. This is a bijective function: any element of the set xH is of the form xh, i.e. $f(h)$, so f is surjective. Now let $f(h_1) = f(h_2)$. Then $xh_1 = xh_2 \implies h_1 = h_2$. Thus, for any $x \in G$, $|xH| = |H|$, and the theorem follows. □

HOMOMORPHISMS OF GROUPS

Just as we did for rings, define a *homomorphism of groups* to be a function $f : G_1 \to G_2$ for groups G_1, G_2, where $f(xy) = f(x)f(y)\ \forall x, y \in G_1$.

It is not necessary to require that $f(\mathbb{1}_{G_1}) = \mathbb{1}_{G_2}$, as this follows from our one condition. As proof, see that

$$f(\mathbb{1}_{G_1}) = f(\mathbb{1}_{G_1})f(\mathbb{1}_{G_1}) \implies \mathbb{1}_{G_2} = [f(\mathbb{1}_{G_1})]^{-1} f(\mathbb{1}_{G_1})f(\mathbb{1}_{G_1}) = f(\mathbb{1}_{G_1})$$

The second consequence of our axiom is that $f(x^{-1}) = [f(x)]^{-1}$. See that

$f(x^{-1})f(x) = f(\mathbb{1}_{G_1}) = \mathbb{1}_{G_2}$, so $f(x^{-1}) = [f(x)]^{-1}$

Similarly, define an *isomorphism of groups* to be a homomorphism $f : G_1 \to G_2$ which is bijective. If there exists such a function, we say $G_1 \cong G_2$, or G_1 is "isomorphic" to G_2. As before, if f is an isomorphism, so is f^{-1}.

Being isomorphic is an equivalence relation on groups. PROPOSITION 6.5

Any two cyclic groups $\langle g_1 \rangle$ and $\langle g_2 \rangle$ are isomorphic IFF $\text{ord}(g_1) = \text{ord}(g_2)$. PROPOSITION 6.6

PROOF. We'll consider the case where both groups are finite. Let $f : \langle g_1 \rangle \to \langle g_2 \rangle$ with $f(g_1^a) = g_2^a$, where $a \in \mathbb{Z}$. This is well defined: let $g_1^a = g_1^b$. Then $g_1^{a-b} = 1$, so $\text{ord}(g_1) =: n | a - b$. Then we say $a - b = kn$, or $a = b + kn$. We have $f(g_1^{b+kn}) = g_2^{b+kn} = g_2^b g_2^{kn} = g_2^b$. However, we also have $f(g_1^{b+kn}) = f(g_1^a) = g_2^a$, so $g_2^a = g_2^b$.

f is also a homomorphism: see that $f(g_1^a g_1^b) = f(g_1^{a+b}) = g_2^{a+b} = g_2^a g_2^b = f(g_1^a)f(g_1^b)$.

Lastly f is bijective: for any element of $\langle g_2 \rangle$, we can write it as g_2^a, which is precisely $f(g_1^a)$. For injectivity, let $f(g_1^a) = f(g_1^b)$. We can write $a = b + kn$, and thus $1 = g_1^{kn} = g_1^a g_1^{-b}$. Multiplying by g_1^b yields $g_1^b = g_1^a$. □

From this, we see that any group with p elements is cyclic and isomorphic to $\mathbb{Z}/p\mathbb{Z}$. As proof, suppose $|G| = p$, and choose $g \neq \mathbb{1} \in G$. Consider $\langle g \rangle < G$. We have $\#\langle g \rangle \| |G| \implies \#\langle g \rangle | p$ by Lagrange. Since $g \neq \mathbb{1}$, $\text{ord}(g) \neq 1$, so $\#\langle g \rangle \neq 1$. Since p is prime, we conclude that $\#\langle g \rangle = p$. Thus, $\langle g \rangle < G$ contain the same elements, and $\langle g \rangle = G \implies G$ is cyclic. Since $\mathbb{Z}/p\mathbb{Z}$ is cyclic with order p, from proposition 6.6, we conclude $\mathbb{Z}/p\mathbb{Z} \cong G$. □

Let $f : G_1 \to G_2$ be a function of any two groups. Define the *kernel* of f to be $\ker(f) := \{g \in G_1 : f(g) = \mathbb{1}_{G_2}\}$. This is identical to our definition for functions between rings.

Let $f : G_1 \to G_2$ be a homomorphism. Then $\ker(f) < G_1$. Furthermore, f is injective IFF $\ker(f) = \mathbb{1}_{G_1}$. PROPOSITION 6.7

PROOF. For the subgroup result, we have: $\mathbb{1}_{G_1} \in \ker(f)$ since $f(\mathbb{1}_{G_1}) = \mathbb{1}_{G_2}$; let $x, y \in \ker(f)$. Then $f(xy) = f(x)f(y) = \mathbb{1}_{G_2}$, so $xy \in \ker(f)$; now let $x \in \ker(f)$. Then $f(x^{-1}) = [f(x)]^{-1} = \mathbb{1}_{G_2}^{-1} = \mathbb{1}_{G_2}$, so $x^{-1} \in \ker(f)$.

For the second claim, suppose $\ker(f) = \mathbb{1}_{G_1}$. Then

$f(x) = f(y) \implies f(x)[f(y)]^{-1} = \mathbb{1}_{G_2} \implies f(xy^{-1}) = \mathbb{1}_{G_2} \implies xy^{-1} \in \ker(f)$.

We assumed $\ker(f) = \mathbb{1}_{G_1}$, so $xy^{-1} = \mathbb{1}_{G_2}$, and finally $x = y$. Conversely, if f is injective, then $f(\mathbb{1}_{G_1}) = \mathbb{1}_{G_2}$ as expected, but this can be the *only* element which maps to $\mathbb{1}_{G_2}$, and we conclude that $\ker(f) = \mathbb{1}_{G_1}$. □

PROPOSITION 6.8 Let $f : G_1 \to G_2$ be a homomorphism and $H < G_1$. Then $f(H)$, i.e. the set $\{f(h) : h \in H\}$, is a subgroup of G_2.

PROOF.

We see immediately that $1_{G_2} \in f(H)$. Let $x, y \in f(H)$. Then $\exists a, b \in H$ with $f(a) = x$ and $f(b) = y$. We have $f(ab) = f(a)f(b) = xy$, so $xy \in f(H)$.

Lastly, we need $x^{-1} \in f(H)$ for $x \in H$. Let $f(a) = x \implies [f(a)]^{-1} = x^{-1}$. By homomorphism properties, $[f(a)]^{-1} = f(a^{-1}) = x^{-1}$. Since H is a subgroup, $a^{-1} \in H$, so $x^{-1} \in f(H)$. □

6.2 Cayley's Theorem
Let G be a finite group with n elements. Then G is isomorphic a subgroup of the symmetric group S_n.

GROUP ACTION ON SETS

Let S be a non-empty set (any set), and let G be a group. We say that G *acts on* S when one defines a function $G \times S \to G : (g, s) \to g * s$ satisfying $\mathbb{1}_G * s = s$ and $g_1 * (g_2 * s) = (g_1 g_2) * s$.

Examples:

1. D_n acts on the vertices of an n-gon, where $x^a(i) = i + a$ and $y(i) = n - i$
2. G acts on itself with $(x, y) \to xyx^{-1}$: we have $\mathbb{1} * y = \mathbb{1}y\mathbb{1} = y$, which is our first condition, and $x_1 * (x_2 * y) = x_1 * (x_2 y x_2^{-1}) = x_1 x_2 y x_2^{-1} x_1^{-1} = (x_1 x_2) * y$
