1. Historical note on algebra

Historically algebra was developed to study zeros of polynomials. The word algebra comes from the name of a book written by a Persian mathematician, Kharazmi (I have to start with this as I am Persian!). In this book, he essentially told us how to find zeros of degree 1 and degree 2 polynomials. Khayyam had a geometric method of solving degree 3 polynomials; the general case was solved by an Italian, Tartaglia. Zeros of degree 4 polynomials were found by Ferrari. In 1824, Abel showed that one cannot express zeros of a general degree 5 polynomial using $+, -, \times, /$, and radicals. In 1832, Galois taught us how to use symmetries of zeros of a polynomial to systematically study them. (You will learn these topics in math100b and math100c!)

Of course, mathematicians, going back to Euclid, have been interested in understanding the symmetries of various objects. In fact, in many cases, symmetries of objects have been our main tool of distinguishing one from others; for instance, in physics, we have to make sure that all the given laws are invariant under the underlying symmetries of the universe or in chemistry symmetries of crystals help us to distinguish them. At the beginning of the 20th century, around the time of discovery of various geometries, e.g. hyperbolic, affine, projective, etc, Klein used symmetries to give a hierarchy of geometries.

Mathematicians observed that a similar structure is needed to be studied in these various examples to get an understanding of symmetries of diverse sets of objects; and that is why group theory was born. Group theory is about understanding symmetries of objects and it is the main topic of this course.

2. Rough idea on what we mean by symmetries.

In preschool we have been ask to find symmetries of shapes like an equilateral triangle. By looking at the triangle, we usually came up with three reflections; we can also notice that there are two rotations; and of course not moving the
shape is another option! Overall an equilateral triangle has 6 symmetries. What do we roughly mean by symmetries of an object \( X \)? The following can be a good rough description: \( \text{Symm}(X) \) consists of functions \( f : X \to X \) such that

1. \( f \) is a bijection;
2. \( f \) preserves properties of \( X \);
3. \( f^{-1} \) preserves properties of \( X \).

(In many examples one can see that the last condition is redundant.) For instance in the example of an equilateral triangle, by properties we mean length and angle. Whatever the set of properties are, it is clear that if \( f_1 \) and \( f_2 \) preserve them, then their composite \( f_1 \circ f_2 \) also does. We also know that the composite of two bijection is again a bijection. Clearly the identity function \( \text{id}_X \) is in \( \text{Symm}(X) \).

Overall we get that \( (\text{Symm}(X), \circ) \) has the following properties:

1. \( \text{id}_X \in \text{Symm}(X) \);
2. \( \forall f_1, f_2 \in \text{Symm}(X), f_1 \circ f_2 \in \text{Symm}(X) \);
3. \( \forall f \in \text{Symm}(X), f^{-1} \in \text{Symm}(X) \).

These are our main motivation to define a **group** as follows.

**Definition 1.** Suppose \( G \) is a non-empty set and \( (g_1, g_2) \mapsto g_1 \cdot g_2 \) is an operation on \( G \); that means it is a function from \( G \times G \) to \( G \). Suppose this operation has the following properties.

1. \( \exists e \in G, \forall g \in G, e \cdot g = g \cdot e = g \) (such an element is called the neutral or the identity element).
2. (Associativity) \( \forall g_1, g_2, g_3 \in G, g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3 \).
3. (inverse) \( \forall g \in G, \exists g' \in G, g \cdot g' = g' \cdot g = e \), where \( e \) is a neutral element.

By the above discussion we have a meta-example of \( (\text{Symm}(X), \circ) \) for any object \( X \).

### 3. Basic examples

1. \( (\mathbb{Z}, +) \) is a group. The neutral element is 0, and inverse of \( x \) with respect to addition is \( -x \):

\[
0 + x = x + 0 = x, \quad x + (-x) = (-x) + x = 0.
\]

2. \( (\mathbb{Z}, \cdot) \) is not a group: the neutral element should be 1; but 2 does not have a multiplicative inverse in \( \mathbb{Z} \).
(3) \((\mathbb{Q},+)\) is a group as in the first example.

(4) \((\mathbb{Q},\cdot)\) is not a group: again the neutral element should be 1; but 0 does not have a multiplicative inverse in \(\mathbb{Q}\).

(5) \((\mathbb{Q}\setminus\{0\},\cdot)\) is a group: 1 is the neutral element and any non-zero element \(\frac{m}{n}\) has a multiplicative inverse \(\frac{n}{m}\).

(6) For \(X = \{1, \ldots, n\}\) with no extra properties, we get all bijections as its group of symmetries; it is denoted by \(S_n\):

\[
S_n := \{f : X \to X | f \text{ is a bijection}\}.
\]

As we discussed above \(S_n\) is a group under composition. Let us recall that \(|S_n| = n!\). We can show this as follows: the number of possibilities for \(f(1)\) is \(n\), after assigning a value to \(f(1)\), there are \(n - 1\) possibilities for \(f(2)\); continuing this logic there are \(n - k + 1\) possibilities for \(f(k)\). Hence we get \(n(n - 1) \cdots (n - k + 1) \cdots 1\) possibilities for a bijection \(f\); and claim follows.

4. Review of basic properties of integers.

We start by reviewing basic properties of integers. We are doing this for two reasons: (1) these properties will be used later to study general groups, (2) these properties motivate us to ask certain questions and wonder to what extent they are true for a general group.

4.1. Well-ordering principle and division algorithm. As you might have seen before, one can construct the rational numbers out of integers. And you will learn in your analysis course, that real numbers can be constructed out of rationals using either Cauchy sequences or Dedekind cuts. For integers, however, we need to use axioms. As Kronecker said “God made the integers; all else is the work of man.” The main axiom that you are familiar with is induction. Induction is equivalent to the well-ordering principle.

Definition 2 (Well-ordering principle). Suppose \(A\) is a non-empty subset of non-negative integers. Then \(A\) has a minimum.

One of the important properties of integers is the division algorithm.
Theorem 3. For any \( a \in \mathbb{Z} \) and \( b \in \mathbb{Z}^+ \), there is a unique pair \((q, r)\) of integers such that (1) \( a = bq + r \), and (2) \( 0 \leq r < b \). (\( q \) is called the quotient and \( r \) is called the remainder.)

In mathematical language we write:

\[
\forall a \in \mathbb{Z}, b \in \mathbb{Z}^+, \exists!(q, r) \in \mathbb{Z} \times \mathbb{Z}, a = bq + r, 0 \leq r < b.
\]

Proof. (Idea of the proof. Going back to elementary school, when we wanted to divide bunch of apples between say 6 people we would have distributed them till we ended up getting less than 6 apples; in each round we were distributing 6 apples and we ended up having fewer than 6 apples; that means if the initial number of apples were \( a \) after \( k \) steps we were ending up with \( a - 6k \) apples. We did this till we get less than 6 apples; this means if we subtract 6 again we get a negative number.)

Existence. Let \( A := \{a - bk \mid k \in \mathbb{Z}, a - bk \in \mathbb{Z} \geq 0\} \) (so \( A = (a + b\mathbb{Z}) \cap \mathbb{Z} \geq 0 \); we denote the set of multiples of \( b \) by \( b\mathbb{Z} := \{bk \mid k \in \mathbb{Z}\} \) and \( a + b\mathbb{Z} \) means its shift by \( a \) that means \( a + b\mathbb{Z} := \{a + bk \mid k \in \mathbb{Z}\} \). So \( A \) is the intersection of \( a + b\mathbb{Z} \) and non-negative integers.) The above argument suggests that the remainder is the minimum of \( A \). By the well-ordering principle, minimum of \( A \) exists exactly when \( A \) is not empty.

Case 1. If \( a \geq 0 \), then \( a = a - b(0) \in A \); and so it is not empty.

Case 2. If \( a < 0 \), then \( a - ab = a(1 - b) \leq 0 \) as \( a < 0 \) and \( 1 - b \leq 0 \); and so \( a - ab \in A \).

Hence in either case, \( A \neq \emptyset \). Since \( A \) is a non-empty subset of non-negative integers, it has a minimum, say \( r \). That implies that \( r = a - bq \) for some integer \( q \) and \( r \geq 0 \). Therefore

\[
a = bq + r, \ 0 \leq r.
\]

To finish the proof of existence, it is remained to show that \( r < b \). Suppose to the contrary that \( r \geq b \). (So it seems we have enough apple for another round!) Then

\[
r - b = (a - bq) - b = a - b(q + 1), \ \text{and} \ r - b \geq 0;
\]

this implies that \( r - b \in A \). Since \( b \) is positive, \( r - b < r \); this contradicts the fact that \( r \) is the minimum element of \( A \). Therefore we get the existence part.

Uniqueness. Suppose \((q_1, r_1)\) and \((q_2, r_2)\) satisfy the desired properties. We have to show \( q_1 = q_2 \) and \( r_1 = r_2 \). We know that \( a = bq_1 + r_1 = bq_2 + r_2 \), which
implies

\[(1) \quad b(q_1 - q_2) = r_2 - r_1.\]

By symmetry without loss of generality we can and will assume that \(q_1 \geq q_2\). Hence by (1) we have \(r_2 - r_1 \geq 0\). Since \(r_1 \geq 0\) and \(r_2 < b\), we get

\[(2) \quad 0 \leq r_2 - r_1 \leq r_2 < b.\]

There is no positive multiple of \(b\) that is less than \(b\); and so by (1) and (2) we have \(r_2 - r_1 = 0\), which means \(r_1 = r_2\). Hence by (1), we get that \(q_1 = q_2\); and claim follows. \qed

4.2. Divisibility. Let us recall that for two integers \(a\) and \(b\) we say \(a\) divides \(b\) and write \(a \mid b\) if there is an integer \(k\) such that \(b = ak\). Here are basic properties of divisibility:

1. If \(a \mid b\) and \(b \mid c\), then \(a \mid c\); alternatively \(a \mid b\) implies \(a \mid bk\) for any integer \(k\).
2. If \(a \mid b_1\) and \(a \mid b_2\), then \(a \mid b_1 \pm b_2\).
3. If \(a \mid b\) and \(b \neq 0\), then \(|a| \leq |b|\).
4. \(a\) is divisible by \(b\) if and only if the remainder of \(a\) divided by \(b\) is 0.

5. Subgroups and subgroup structure of \(\mathbb{Z}\)

Whenever you learn about a new structure in mathematics, you should ask about its substructures and the maps that preserve that structure. For instance, in linear algebra, after learning vector spaces, you study subspaces and linear maps. In group theory the substructures are called subgroups and the maps that preserve their structure are called homomorphisms. We will come back to these topics later; but for now let us just say what a subgroup is.

**Definition 4.** Suppose \((G, \cdot)\) is a group. We say a subset \(H\) of \(G\) is a subgroup if \(H\) is a group under the same operation \(\cdot\).

Notice that by definition, if \(H\) is a subgroup of \((G, \cdot)\), then for any \(h_1\) and \(h_2\) in \(H\) we have \(h_1 \cdot h_2 \in H\). (We say \(H\) is closed under the operation \(\cdot\).) The converse is not necessarily true: consider the group \((\mathbb{Z}, +)\). Non-negative integers \(\mathbb{Z}^{\geq 0}\) is closed under addition; that means sum of two non-negative integers is a non-negative integer; but \(\mathbb{Z}^{\geq 0}\) is not a subgroup: for instance the additive inverse of 2 which is \(-2\) is not in \(\mathbb{Z}^{\geq 0}\).
The set of even numbers is a subgroup of \( \mathbb{Z} \); more generally for any \( a \in \mathbb{Z} \), the set \( a\mathbb{Z} := \{ka | k \in \mathbb{Z}\} \) of all the multiplies of \( a \) is a subgroup of \( \mathbb{Z} \). Here is why:

- Clearly + is associative, so it is enough to show the following:
  1. \( 0 \in a\mathbb{Z} \),
  2. If \( b_1, b_2 \) are in \( a\mathbb{Z} \), then \( b_1 + b_2 \) is in \( a\mathbb{Z} \).
  3. If \( b \) is in \( a\mathbb{Z} \), then \( -b \) is in \( a\mathbb{Z} \).

Notice that by the definition of divisibility \( b \in a\mathbb{Z} \) if and only if \( a | b \). \( 0 = (a)(0) \) implies that \( 0 \in a\mathbb{Z} \).

We have \( b_1, b_2 \in a\mathbb{Z} \implies a|b_1 \) and \( a|b_2 \implies a|b_1 + b_2 \implies b_1 + b_2 \in a\mathbb{Z} \).

We also have \( b \in a\mathbb{Z} \implies a|b \implies a|0 - b \) (as \( a|0 \)) \implies \( a|(-b) \implies -b \in a\mathbb{Z} \).

Next we show any subgroup of \( \mathbb{Z} \) is of this form.

**Theorem 5** (Subgroup structure of \( \mathbb{Z} \)). \( H \) is a subgroup of \( \mathbb{Z} \) if and only if \( H = a\mathbb{Z} \) for some \( a \in \mathbb{Z} \).

**Proof.** We have already proved this (\( \Leftarrow \)) direction. So we focus on this (\( \Rightarrow \)) direction.

If \( H = \{0\} \), then \( a = 0 \) works; and we are done. So we can and will assume there is \( h \in H \setminus \{0\} \). Since \( H \) is a subgroup, \( -h \in H \). Since \( h \neq 0 \), either \( h \) or \( -h \) is positive. Overall we get that \( H \cap \mathbb{Z}^+ \neq \emptyset \); that means there is a positive element in \( H \). Hence by the well-ordering principle, \( H \cap \mathbb{Z}^+ \) has a minimum. Suppose \( a \in H \) is the minimum positive element of \( H \). We want to show \( H = a\mathbb{Z} \).

**Step 1.** \( a\mathbb{Z} \subseteq H \). We have to show for any integer \( k \), \( ak \in H \). We first use induction to prove this for non-negative integers.

- **Base of induction.** \( (a)(0) = 0 \in H \) as \( H \) is a subgroup.

- **The induction step.** We have to show that if \( ak \in H \), then \( a(k + 1) \in H \).

  By the induction hypothesis, \( ak \in H \). Since \( a \in H \) and \( H \) is closed under addition, we get that \( a + ak \in H \), which means \( a(k + 1) \in H \); and claim follows.

  If \( k \) is a negative integer, then \( -k \) is a positive integer. Hence by the above discussion \( a(-k) \in H \), which means \( -ak \in H \). Since \( H \) is a subgroup, we get that \( ak = -(-ak) \in H \).
Overall we get that for any integer $k$ we have $ak \in H$; this means $a\mathbb{Z} \subseteq H$.

**Step 2.** $H \subseteq a\mathbb{Z}$. For any $h \in H$, we have to show $h$ is a multiple of $a$; that means we have show the remainder of $h$ divided by $a$ is zero. Suppose $q$ is the quotient and $r$ is the remainder of $h$ divided by $a$. Then $h = aq + r$ and $0 \leq r < a$. Hence $r = h - aq$. By Step 1 we know that $-aq \in H$. Since $h \in H$ and $H$ is closed under addition, $r = h - aq \in H$. Since $r \in H$, $r < a$, and $a$ is the smallest positive number in $H$, we deduce that $r$ cannot be positive; that means $r \leq 0$. On the other hand, we know that $r$ is not negative, $r \geq 0$. Hence overall $r = 0$. This means $h = aq \in a\mathbb{Z}$; and claim follows.

By Steps 1 and 2, we get that $H = a\mathbb{Z}$. \hfill \Box

6. **The greatest common divisor of two integers.**

Let us recall that the greatest common divisor of $a$ and $b$ is denoted by $\gcd(a, b)$ and by the definition, it is the largest number that divides both $a$ and $b$.

Notice that $\gcd(0, 0)$ does not make sense. Also observe that $\gcd(a, b)$ is at most $\min(|a|, |b|)$ if $a$ and $b$ are not zero.

We will study the basic properties of the greatest common divisors in the next lecture.