Notations: we write \( \mathbb{Z}^+ = \{0, 1, 2, \ldots\}, \mathbb{N} = \{1, 2, 3, \ldots\}, 2\mathbb{N} := \{2n; n \in \mathbb{N}\} = \{2, 4, 6, \ldots\}, \) and \( 2\mathbb{N} - 1 := \{2n - 1; n \in \mathbb{N}\} = \{1, 3, 5, \ldots\} \).

Semigroups

In the following, \( \circ : S \times S \to S \) is a binary operation on a nonempty set \( S \). We write \( a \circ b \) instead of the too formal \( \circ(a, b) \).

We say that \( \circ \) is **associative** if \( (a \circ b) \circ c = a \circ (b \circ c) \) for all \( a, b, c \in S \). This is the most important property, and all our operations below will share that.

**Definition.** Let \( \circ \) be an associative binary operation on a nonempty set \( S \). Then the pair \((S, \circ)\) is a called a **semigroup**. (When the operation \( \circ \) is clear from the context, we often just say that \( S \) is a semigroup.)

**Remark.** Implicit in the word “binary operation” is the following property - often called closure: for all \( a, b \in S \), \( a \circ b \in S \).

We say that \( \circ \) is **commutative** if \( a \circ b = b \circ a \) for all \( a, b \in S \). [Warning: our operations are not assumed to be commutative unless explicitly stated so!]

We say that \((S, \circ)\) has an **identity** (or “\( S \) has an identity” for short) if there is an element \( e \in S \) such that \( e \circ x = x \) and \( x \circ e = x \) for all \( x \in S \). An identity is also called a neutral element. It is easy to see that when exists, the identity is unique. [Indeed, if \( e \) and \( e' \) are identities, then \( e = e \circ e' = e' \).] A semigroup with identity is sometimes called a monoid.

When a semigroup \((S, \circ)\) has an identity \( e \), we say that an element \( a \in S \) has an **inverse** (or “\( a \) is invertible”, or “\( a \) is a unit”) if there is a \( b \in S \) such that \( a \circ b = e \) and \( b \circ a = e \); it is easy to see that if exists, such an element \( b \) is unique; we usually write \( a^{-1} \) for this \( b \) and call it the inverse of \( a \). [Indeed, if \( b \) and \( b' \) are two such elements, then \( b = b \circ (a \circ b') = (b \circ a) \circ b' = b' \).] The set of all invertible elements of \( S \) is denoted by \( S^* \).

**Examples.** \((\mathbb{R}, +)\), \((\mathbb{Z}, +)\), \((\mathbb{Z}^+, +)\), \((\mathbb{N}, \cdot)\), \((2\mathbb{N} - 1, \cdot)\), \((\mathbb{Z}_n, +)\), \((\mathbb{Z}_n, \cdot)\), as well as the set of \( n \times n \) real matrices with respect to matrix-multiplication are semigroups with identity. \((\mathbb{N}, +)\) and \((2\mathbb{N}, \cdot)\) are semigroups without identity.

Groups

**Definition.** Let \( \circ \) be an associative binary operation on a nonempty set \( G \). The pair \((G, \circ)\) is a called a **group** if \( G \) has an identity and each element of \( G \) has an inverse. The number of elements in \( G \) is called the **order** of the group. When \( \circ \) is commutative, we say that the group \((G, \circ)\) is commutative or **Abelian**.

**Remark.** We often use \( \cdot \) to denote the group operation and call it multiplication. Then we may just write \( ab \) for \( a \cdot b \), and sometimes we write \( 1 \) to denote the identity. For commutative groups, we often use \( + \) to denote the operation and call it addition, write \( 0 \) for the identity, and \((-a)\) for the inverse of \( a \).
Theorem 1. Let \((S, \circ)\) be a semigroup with identity (a monoid). Then \((S^*, \circ)\) is a group.

Examples. Of the above semigroup examples, the only ones that are groups are \((\mathbb{R}, +)\), \((\mathbb{Z}, +)\) and \((\mathbb{Z}_n, +)\). Here are a few more standard Abelian groups: \((\mathbb{Q}, +)\), \((\mathbb{C}, +)\), \((\mathbb{Q}\setminus\{0\}, \cdot)\), \((\mathbb{R} \setminus \{0\}, \cdot)\), \((\mathbb{C} \setminus \{0\}, \cdot)\). The set of all \(k \times k\) non-singular real matrices forms a non-Abelian group with respect to matrix-multiplication; so does the set of all symmetries of an equilateral triangle (discussed in class) under composition.

A useful (counter)example: Let \(S\) be a set containing at least two elements. Define a binary operation \(\cdot\) on \(S\) by \((\forall x, y \in S)x \cdot y = x\). How do the group axioms fare for \(S\) equipped with this operation? Firstly, \(\cdot\) is clearly associative. Furthermore, the condition \(|S| \geq 2\) implies that \(S\) has no left-identity (hence no identity), but every element of \(S\) is a right-identity. This example may be useful for discarding some hastily made conjectures about groups. One could even add an identity \(e\) to \(S\) and still keep its weirdness.

Subgroups

Definition. Let \((G, \circ)\) be a group. A subset \(S\) of \(G\) is a subgroup if \(S\) itself is a group with respect to the same operation \(\circ\). We write \((S, \circ) \leq (G, \circ)\), or simply write \(S \leq G\) when it is clear what the operation is. \(S < G\) means \(S \leq G\) and \(S \neq G\) (proper subgroup).

It is easy to see that a nonempty subset of \(G\) forms a subgroup with respect to \(\circ\) if and only if it is closed under \(\circ\) and is closed under taking inverse (in \((G, \circ)\)). The following test combines these two into one:

Theorem 2 (Closure Test). Let \((G, \circ)\) be a group and let \(S \subseteq G\) be nonempty. Then \((S, \circ)\) is a group if and only if \(a \circ b^{-1} \in S\) for all \(a, b \in S\).

Examples: \((2\mathbb{Z}, +) < (\mathbb{Z}, +) < (\mathbb{Q}, +) < (\mathbb{R}, +) < (\mathbb{C}, +)\) and \((\mathbb{Q}^*, \cdot) < (\mathbb{R}^*, \cdot) < (\mathbb{C}^*, \cdot)\) are subgroup relations.

Theorem 3 (\(\mathbb{Z}\)). The only subgroups of \((\mathbb{Z}, +)\) are the sets \(d\mathbb{Z} := \{dn : n \in \mathbb{Z}\}\), \(d = 0, 1, 2, \ldots\) [Hint for a proof: let \(I \leq \mathbb{Z}\) and start with the smallest positive element of \(I\) (if any).]

Corollary. Let \((G, \cdot)\) be a group with identity \(e\), and let \(a \in G\) be arbitrary. The set \(\{k \in \mathbb{Z} : a^k = e\}\) is clearly a subgroup of \(\mathbb{Z}\), and hence it is of the form \(d\mathbb{Z}\) for some nonnegative integer \(d\). When this \(d\) is positive, we say that the order of \(a\) is \(d\), and we write \(o(a) = d\). Thus, the order of \(a\) is the smallest positive integer \(d\) (if any) such that \(a^d = e\).

Theorem 4 (Lagrange). Let \(G\) be a finite group of order \(n\) with identity \(e\). Then, \(a^n = e\) for all \(a \in G\). Hence, the order of any element of \(G\) is a divisor of \(n\). More generally, the order of any subgroup of \(G\) divides \(n\).

Remark. One can get an easy proof for commutative groups by using the following lemma. (For non-commutative groups the standard proofs use the notion of cosets.)

Lemma. Let \((G, \circ)\) be a group and let \(a \in G\) be arbitrary. The map \(f_a : G \to G : x \mapsto a \circ x\) is a bijection.

Proof of Lagrange's theorem in the commutative case: Let \(a \in G\). By the previous lemma,

\[
\prod_{g \in G} g = \prod_{g \in G} (ag) = a^{\lvert G \rvert} \prod_{g \in G} g
\]

and the claim follows. \(\square\)
Some number theory

We will show now how to obtain the Fundamental Theorem of Arithmetic based purely on Euclid’s 2300 years old ingenious invention: the Euclidean Algorithm. One advantage of this approach is that it generalizes to similar algebraic structures, e.g., to the ring of polynomials.

**Theorem 5.** Given \(a, b \in \mathbb{Z}\), not both 0, there exists a (unique) positive integer \(d\) such that \(d\) is a common divisor of \(a\) and \(b\) [divides both \(a\) and \(b\)], and if \(k\) is any common divisor of \(a\) and \(b\), then \(k|d\). This number \(d\) is called the greatest common divisor of \(a\) and \(b\) [since it happens to be the same as the largest one of all common divisors], and it is denoted by \(\gcd(a, b)\). The greatest common divisor of two numbers is computed by – and hence its existence is proved by – the Euclidean Algorithm; see http://en.wikipedia.org/wiki/Euclidean_algorithm

**Theorem 6 (Integer Division Theorem).** For every \(a \in \mathbb{Z}\) and \(b \in \mathbb{N}\) there are \(q, r \in \mathbb{Z}\) such that \(a = qb + r\) and \(0 \leq r < b\).

**Theorem 7 (GCD Theorem).** Let \(a\) and \(b\) be non-zero integers. Then there are integers \(x\) and \(y\) such that \(\gcd(a, b) = ax + by\). In fact, writing \(d = \gcd(a, b)\), we have

\[
\{ax + by : x, y \in \mathbb{Z}\} = d\mathbb{Z} := \{dn : n \in \mathbb{Z}\}.
\]

**Remark.** The Extended Euclidean Algorithm — see www.millersv.edu/~bikenaga/absalg/exteuc/exteucth.html — computes one such pair \((x, y)\) (as well as \(\gcd(a, b)\)), yet we give a direct proof below to Theorem 7 (which would thus also prove Theorem 5).

**Proof.** Let \(s\) be the smallest positive member of the set \(S := \{ax + by : x, y \in \mathbb{Z}\}\). We will show that \(s = d\). (The claim \(\{ax + by : x, y \in \mathbb{Z}\} = d\mathbb{Z}\) then easily follows.)

Now, \(d\) obviously divides all elements of the set \(S\), hence \(d|s\) and thus \(s \geq d\). We show next that \(s|a\) and \(s|b\), that is, \(s\) a common divisor of \(a\) and \(b\) and thus \(s \leq d\) (\(d\) being the greatest common divisor). Indeed, apply the Integer Division Theorem to \(a\) and \(s\) to find \(q\) and \(r\) such that \(a = qs + r\) and \(0 \leq r < s\). Since \(r = a − qs\) and \(s\) is of the form \(ax + by\) \((x, y \in \mathbb{Z})\), so \(r\) is also of this form. But then \(0 \leq r < s\) implies \(r = 0\) (since \(s\) was the smallest positive number of this form). The proof of \(s|b\) is similar.

**Corollary.** The (Diophantine) equation \(ax + by = c\) has a solution (in integers \(x, y\)) if and only if \(\gcd(a, b)\) divides \(c\).

In other words, the congruence \(ax \equiv c \pmod{m}\) has a solution \(x\) if and only if \(\gcd(a, m)\) divides \(c\); and in that case there are exactly \(\gcd(a, m)\) different solutions modulo \(m\).

In particular (setting \(c = 1\) above), \(a\) has a multiplicative inverse modulo \(m\) if and only if \(\gcd(a, m) = 1\).

**Theorem 8.** If \(a\) divides \(b \cdot c\), and \(a\) and \(b\) are relatively prime, then \(a\) divides \(c\).

**Proof.** By the GCD Theorem, there are \(x, y\) such that \(1 = \gcd(a, b) = ax + by\). Hence \(c = acx + bcy\), and since both \(acx\) and \(bcy\) are divisible by \(a\), so is \(c\).

**Corollary.** If a prime \(p\) divides \(b \cdot c\), then either \(p\) divides \(b\) or \(p\) divides \(c\).

**Corollary (The Fundamental Theorem of Arithmetic).** Any integer greater than 1 can be factored uniquely as a product of primes.
A simple application

**Theorem 9.** Let \( k, n \in \mathbb{N} \). Then \( \sqrt[n]{k} \) is either integer or irrational.

**Proof.** Assume \( \sqrt[n]{k} \) is rational, say \( p/q \) where \( p, q \in \mathbb{N} \), and \( \gcd(p, q) = 1 \) (simplify the fraction otherwise). We need to show that \( q = 1 \).

Now, \( nq^k = p^k \). Thus \( q \) divides \( p^k = p \cdot p^{k-1} \), and hence, by Theorem 8, \( q \) divides \( p^{k-1} \).

Applying (inductively) this argument \( k \) times shows that \( q \) divides 1, hence \( q = 1 \). \( \square \)

**Corollaries of Lagrange’s Theorem**

**Theorem 10 (Fermat’s Little Theorem).** Let \( p \) be prime and \( a \in \mathbb{Z} \) such that \( p \nmid a \). Then,

\[
a^{p-1} \equiv 1 \pmod{p}.
\]

This theorem is a special case of Euler’s theorem (see below).

**Definition.** For \( m \in \mathbb{N} \), we define the Euler (totient) function \( \varphi(m) \) as follows: \( \varphi(m) \) is the number of integers between 1 and \( m \) that are relatively prime to \( m \):

\[
\varphi(m) := |\{k : 1 \leq k < m, \gcd(k, m) = 1\}|.
\]

**Theorem 11 (Euler’s Theorem).** Let \( m \in \mathbb{N}, m \geq 2 \), and let \( a \) be relatively prime to \( m \). Then,

\[
a^{\varphi(m)} \equiv 1 \pmod{m}.
\]

**Proof.** Indeed, the set \( S := \{k : 1 \leq k < m, \gcd(k, m) = 1\} = \mathbb{Z}_m^* \) (the set of invertible elements of \( \mathbb{Z}_m \)) forms a group under multiplication modulo \( m \). Hence the claim follows from Lagrange’s theorem (which we proved in the commutative case). \( \square \)

**Remark.** It is not hard to find the following explicit formula for \( \varphi(m) \): If \( m = p_1^{\alpha_1}p_2^{\alpha_2} \cdots p_k^{\alpha_k} \) where \( p_i \) are distinct primes, then

\[
\varphi(m) = m \prod_{i=1}^{k} \left( 1 - \frac{1}{p_i} \right).
\]

In particular, if \( n = pq \) where \( p \) and \( q \) are distinct primes, then \( \varphi(n) = (p-1)(q-1) \). (This is used in the RSA scheme.)