3.1.10 An Invertible Element Can Be Canceled.

Proposition. Let $(S, \circ, e)$ be a monoid, and let $a \in S$ is its invertible element. Then

$$a \circ b = a \circ c, \text{ or } b \circ a = c \circ a \quad \text{implies} \quad b = c.$$

□

Justification. Assume that a^{-1} exists and

$$a \circ b = a \circ c. \tag{3.4}$$

Multiply 3.4 by a^{-1} form the left. We get

$$a^{-1} \circ (a \circ b) = a^{-1} \circ (a \circ c), \text{ which gives } (a^{-1} \circ a) \circ b = (a^{-1} \circ a) \circ c \text{ and } b = c.$$

Similarly for $b \circ a = c \circ a$. The only difference is that here we multiply by a^{-1} from the right. (Notice the similarity with matrix operations.) □

3.1.11 Groups. In couple of examples above, every element was invertible; indeed, it holds for $(\mathbb{Z}, +, 0)$, $(\mathbb{R} \setminus \{0\}, \cdot, 1)$, and $(\mathbb{Z}_n, +, 0)$. Such monoids are of great importance and they are called groups.

Definition. A monoid $(S, \circ, e)$ in which every element is invertible is called a *group*. □

Examples of groups. The following monoids are groups:

1) The monoid $(\mathbb{R}, +, 0)$. Indeed, for every $x \in \mathbb{R}$ there exists $-x$ for which $x + (-x) = 0 = (-x) + x$.
2) The monoid $(\mathbb{Z}, +, 0)$. Indeed, for each integer x there exists an integer $-x$ for which $x + (-x) = 0 = (-x) + x$.
3) The monoid $(\mathbb{R}^+, \cdot, 1)$, where $\mathbb{R}^+$ is the set of all positive real numbers. Indeed, for every positive real number x there exists a positive real number $\frac{1}{x}$ for which $x \cdot \frac{1}{x} = 1 = \frac{1}{x} \cdot x$.
4) The monoid $(\mathbb{Z}_n, \oplus, [0]_n)$. Indeed, for a class $[i]_n$ there exists a class $[n - i]_n$ for which $[i]_n \oplus [n - i]_n = [0]_n = [n - i]_n \oplus [i]_n$.
5) Let A be the set of all permutation of the set $\{1, 2, \ldots, n\}$, and let $\circ$ be the composition of permutations. Then $(A, \circ)$ is a monoid with the neutral element the identity permutation id. Moreover, for every permutation ϕ there exists its inverse permutation ϕ^{-1} for which $\phi \circ \phi^{-1} = id = \phi^{-1} \circ \phi$.

Examples of monoids that are not groups.

1) The monoid $(\mathbb{Z}, \cdot, 1)$. Indeed, for example 2 is not invertible because there is no **integer** k such that $2 \cdot k = 1$.
2) The monoid $(\mathbb{Z}_n, \odot, [1]_n)$. Indeed, the class $[0]_n$ is not invertible because for any $[i]_n$ we have $[0]_n \odot [i]_n = [0]_n \neq [1]_n$.
3) Let B be the set of all mappings from the set $\{1, 2, \ldots, n\}$ into itself, where $n > 1$. Let $\circ$ be the composition of mappings. Then $(B, \circ, id)$ is a monoid where id is the identity mapping. Any mapping that is not one-to-one is not invertible.

3.1.12 Groups can be characterized as those semigroups $(S, \circ)$ where every equation $a \circ x = b$ and $y \circ a = b$ has a solution. In that case, the solution is unique. From this it immediately follows that

1. If $(S, \circ)$ is not a group, then there is an equation which does not have a solution.
2. Given a semigroup $(S, \circ)$. If there exists an equation with two distinct solutions, then $(S, \circ)$ is not a group, and moreover there is an equation that does not have a solution.

The following two paragraphs prove it.

3.1.13 Proposition. Given a group $(S, \circ)$ with its neutral element e. Then for every two elements $a, b \in S$ there exist unique $x, y \in S$ such that

$$a \circ x = b, \qquad y \circ a = b.$$

□

Justification. Since $(S, \circ, e)$ is a group and $a \in S$, there exists its inverse a^{-1}. If we multiply the equation $a \circ x = b$ by a^{-1} from the left we obtain

$$x = (a^{-1} \circ a) \circ x = a^{-1} \circ (a \circ x) = a^{-1} \circ b.$$

Similarly we obtain $y = b \circ a^{-1}$ from the second equation; indeed, we multiply the second equation by a^{-1} from the right and get the desired solution.

Let us show the uniqueness. Assume that $a \circ x_1 = b$ and $a \circ x_2 = b$. Then $a \circ x_1 = a \circ x_2$. Now, the proposition 3.1.10 completes the argument because it states that $x_1 = x_2$. Similarly from $y_1 \circ a = b$ and $y_2 \circ a = b$ we get $y_1 = y_2$.

3.1.14 Theorem. A semigroup $(S, \circ)$ is a group if and only if every equation of the form $a \circ x = b$ and every equation of the form $y \circ a = b$ has at least one solution.

More precisely: A semigroup $(S, \circ)$ is a group if and only if for every two elements $a, b \in S$ there exist $x, y \in S$ such that $a \circ x = b$ and $y \circ a = b$. □

Justification. First we show that if a semigroup $(S, \circ)$ satisfies the above conditions then it has got a neutral element.

Choose any $a \in S$. There exists $e_a \in S$ such that $e_a \circ a = a$; indeed, it is a solution of $y \circ a = a$. Now, take an arbitrary $b \in S$. We know that $b = a \circ x$ for some $x \in S$, hence

$$e_a \circ b = e_a \circ (a \circ x) = (e_a \circ a) \circ x = a \circ x = b.$$

Similarly, it can be shown that the element f_a for which $a \circ f_a = a$ satisfies $b \circ f_a = b$ for any $b \in S$.

Therefore, from 3.1.4 we get that $e_a = f_a$ is the neutral element of $(S, \cdot)$.

To show that every element $a \in S$ is invertible, it suffices to use the proposition from 3.1.7. Indeed, from the fact that there exist $x, y \in S$ with $a \circ x = e$ and $y \circ a = e$ we know that $x = y$, and $x = a^{-1}$. So, a is invertible. Since a was an arbitrary element of S, $(S, \circ, e)$ is a group. □

3.1.15 Commutative Semigroups, Monoids, Groups. In many examples above (but not in all) it does not matter whether we calculate $a \circ b$ or $b \circ a$, we get the same results.

Definition. A semigroup $(S, \circ)$ (monoid, group) is called *commutative* if it satisfies the *commutative law*, i.e. for every two elements $x, y \in S$

$$x \circ y = y \circ x.$$

□

3.1.16 Subsemigroups. Given a semigroup $(S, \circ)$ and a set $T \subseteq S$. It may happen (but does not need to) that T together with the same operation $\circ$ is again a semigroup. In that case, we will call $(T, \circ)$ a subsemigroup of $(S, \circ)$.

Definition. Given a semigroup $(S, \circ)$. A subset $T \subseteq S$ together with an operation $\circ$ forms a *subsemigroup* of the semigroup $(S, \circ)$, if for every two elements $x, y \in T$ we have $x \circ y \in T$. (In this case $(T, \circ)$ is also a semigroup.) □

Remark. Next, we will say less exactly "T is a subsemigroup" instead of "T forms a subsemigroup". It will be mainly in the situation where the operation is clear from the context.

Examples of subsemigroups. The following are examples of subsemigroups:

1) $\mathbb{N}$ together with addition forms a subsemigroup of $(\mathbb{Z}, +)$.
2) The set of all regular matrices together with multiplication of matrices forms a subsemigroup of $(M_n, \cdot)$, where M_n is the set of all square matrices of order n.
3) The set of all positive real numbers together with multiplication forms a subsemigroup of $(\mathbb{R}, \cdot)$.

Example of a subset that does not form a subsemigroup. The set of all regular square matrices of order n together with addition of matrices does not form a subsemigroup of $(M_n, +)$. Indeed, it does not hold that sum of two regular matrices is a regular matrix, e.g. coincide the identity matrix E. Then E and $-E$ are regular matrices but $E + (-E)$ is the zero matrix which is not regular.

3.1.17 Submonoids.

Definition. Given a monoid $(S, \circ, e)$. A subset $T \subseteq S$ forms a submonoid if it forms a subsemigroup and moreover $e \in T$. (In this case $(T, \circ, e)$ is also a monoid.) □

Examples of submonoids.

1) The set of all natural numbers $\mathbb{N}$ together with addition is a submonoid of $(\mathbb{Z}, +, 0)$, since $0 \in \mathbb{N}$.
2) The set of all regular square matrices of order n together with multiplication of matrices forms a submonoid of $(M_n, \cdot, E)$, since the identity matrix E is regular.
3) Denote by T_X the set of all mappings from a set X into itself. Consider the operation composition of mappings $\circ$. Then $(T_X, \circ, id)$ where id is the identity mapping (defined by $id(x) = x$ for all $x \in X$) is a monoid. The set of all bijections from T_X forms a submonoid of $(T_X, \circ)$, indeed, a composition of two bijections is a bijection, and the identity mapping is a bijection.

3.1.18 Remark. Notice that a subsemigroup $(T, \circ)$ of $(S, \circ, e)$ may contain a neutral element which is different from the neutral element e (but in this case $e \notin T$). If this is the case $(T, \circ)$ is a subsemigroup of $(S, \circ)$ but not a submonoid of $(S, \circ, e)$. Next, there is an example of such a situation.

Example. Let $X = \{1, 2, 3\}$. Denote by S the set of all mappings from X to X. Then $(S, \circ, id)$ is a monoid ($\circ$ is the composition of mappings, id is the identity mapping).

Consider the mapping $f\colon X \to X$ defined by $f(1) = 2$, $f(2) = 3$, $f(3) = 4$, and $f(4) = 2$. Then $f^4 = f$ and $T = \{f, f^2, f^3\}$ forms a subsemigroup of $(S, \circ, id)$. T does not form a submonoid, since $id \notin T$. On the other hand, f^3 is the neutral element of $(T, \circ)$ and $(T, \circ, f^3)$ is in fact a group. Indeed, $f \circ f^3 = f = f^3 \circ f$, $f^2 \circ f^3 = f^2 = f^3 \circ f^2$, and $f^3 \circ f^3 = f^3$.

3.1.19 The Group of Invertible Elements. Every monoid contains a special submonoid, the one formed by all invertible elements. And this submonoid is in fact a group that is called the *group of invertible elements*. Let us first prove the following proposition which justifies the definition coming next.

Proposition. Given a monoid $(S, \circ, e)$. Denote by $S^\star$ the set of all its invertible elements. Then $(S^\star, \circ, e)$ is a submonoid of $(S, \circ)$ which is a group. □

Justification. The above proposition immediately follows from 3.1.9. Indeed, $e \in S^\star$, and if $a, b \in S^\star$ then $a \circ b \in S^\star$. So $S^\star$ forms a submonoid.

Moreover, $(S^\star, \circ, e)$ is a group because if $a \in S^\star$ then $a^{-1} \in S^\star$. □

Definition. The group $(S^\star, \circ, e)$ is called the *group of invertible elements* of the monoid S. □

3.1.20 The following theorem is an important fact and is used in a lot of applications. In fact it holds for any finite group but we will state and prove it only for commutative ones now.

Theorem. Let $(G, \circ, e)$ be a finite commutative group. Then for every $a \in G$ we have $a^{|G|} = e$. □

Justification. Assume that the group has n elements and denote $G = \{a_1, a_2, \ldots, a_n\}$. Take any $a \in G$ and form the set $H = \{a \circ a_1, a \circ a_2, \ldots, a \circ a_n\}$. The H has also n elements; indeed, if $a \circ a_i = a \circ a_j$ in a group then $a_i = a_j$ (see 3.1.10).

Therefore, $G = H$ and because G is a commutative group we have

$$a_1 \circ a_2 \circ \ldots \circ a_n = (a \circ a_1) \circ (a \circ a_2) \circ \ldots \circ (a \circ a_n),$$

and also

$$a_1 \circ a_2 \circ \ldots \circ a_n = a^n \circ (a_1 \circ a_2 \circ \ldots \circ a_n).$$

If we multiply the last equality by $(a_1 \circ a_2 \circ \ldots \circ a_n)^{-1}$ we get $a^n = e$. □

3.2 Applications to $(\mathbb{Z}_n, \cdot, 1)$

Let us first introduce the *Euler function.*

3.2.1 Euler function. Given a natural number $n > 1$. Then the value of Euler function $\phi(n)$ equals to the number of all natural numbers i, $0 \leq i < n$, that are relatively prime to n. □

For example $\phi(6) = 2$, since there are only two natural numbers between 0 and 5 that are relatively prime to 6, namely 1 and 5.

3.2.2 Properties of Euler Function.

1. Let p be a prime number, then $\phi(p) = p - 1$.
2. If p is a prime number and $k \geq 1$ then $\phi(p^k) = p^k - p^{k-1}$.
3. If n and m are relatively prime natural numbers then $\phi(n \cdot m) = \phi(n) \cdot \phi(m)$.

□

It is not difficult to show the first two properties above. The easiest way how to prove the last one is to use the Chinese Remainder Theorem which is beyond the scope of this course.

3.2.3 The Group of Invertible Elements of $(\mathbb{Z}_n, \cdot, 1)$. We will use the facts from 3.1.19 for the commutative monoid $(\mathbb{Z}_n, \cdot, 1)$. We know $(\mathbb{Z}_n, \cdot)$ is a monoid with its neutral element 1. The set of all invertible elements of it is

$$\mathbb{Z}_n^\star = \{i \,|\, 0 \leq i < n, \;\; i \text{ and } n \text{ are relatively prime}\}.$$

Therefore, $(\mathbb{Z}_n^\star, \cdot, 1)$ is a group with $\phi(n)$ elements where $\phi(n)$ is the Euler function of n.

3.2.4 Euler-Fermat Theorem. Applying 3.1.20 we get a theorem which generalizes of the small Fermat theorem:

Theorem (Euler-Fermat). Given a natural number $n > 1$. Then for every integer a relatively prime to n we have

$$a^{\phi(n)} \equiv 1 \,(\mathrm{mod}\, n).$$

□

Justification. Indeed, take any integer a relatively prime to n. Put b to be the remainder when we divide a by n. Then $b \in \mathbb{Z}_n^\star$. Since $(\mathbb{Z}_n^\star, \cdot, 1)$ is a finite group with $\phi(n)$ elements, the Euler-Fermat Theorem is a consequence of 3.1.20. □

Remark. The small Fermat theorem is an immediate consequence of the Euler-Fermat theorem. Indeed, if n is a prime number then $\phi(n) = n - 1$.

3.3 Subgroups

Analogously as we defined subsemigroups and submonoids we can define subgroups. Subgroups are formed by subsets that not only form itself a group but group with the original operations. More precisely:

Definition. Given a group $(G, \circ, e)$. We say that $H \subseteq G$ forms a *subgroup* of $(G, \circ, e)$ if

1. for every $x, y \in H$ it holds that $x \circ y \in H$, (i.e. forms a subsemigroup);
2. $e \in H$, (i.e. forms a submonoid);
3. for every $x \in H$ it holds that $x^{-1} \in H$.

□

Note, that in this case, $(H, \circ, e)$ is also a group.

Remark. Every group $(G, \circ, e)$ with more than one element has at least two subgroups; indeed, one formed by $\{e\}$ and second formed by G. These two subgroups are called *trivial subgroups*.

3.3.1 How Many Elements a Subgroup Can Have? We will show some useful properties of finite groups and their subgroups. The first theorem shows that a subset of a group can form a subgroup only if its number of elements divides the number of elements of the group. Hence, $(\mathbb{Z}_7, +, 0)$ has only trivial subgroups; indeed, 7 is a prime number with divisors 1 and 7. And any subgroup with 1 element consists of 0, a subgroup with 7 elements is $(\mathbb{Z}_7, +, 0)$.

Theorem. Let $(G, \circ, e)$ be a finite group and $H \subseteq G$ its subgroup. Then the number of elements of H divides the number of elements of G. □

Justification. Let us denote $n = |G|$ and $k = |H|$. For every $g \in G$ we form a subset of G: $g \circ H = \{g \circ x \,|\, x \in H\}$.

We show that for every $g_1, g_2 \in G$ the sets $g_1 \circ H$ and $g_2 \circ H$ are either the same or they are disjoint (they do not have a common element).

Assume that $(g_1 \circ H) \cap (g_2 \circ H) \neq \emptyset$. Then there exist $h_1, h_2 \in H$ such that $g_1 \circ h_1 = g_2 \circ h_2$. Since we are in a group, we have

$$g_1 = (g_2 \circ h_2) \circ h_1^{-1} = g_2 \circ (h_2 \circ h_1^{-1}) \text{ and } g_2 = (g_1 \circ h_1) \circ h_2^{-1} = g_1 \circ (h_1 \circ h_2^{-1}). \quad (3.5)$$

This means that $g_1 \in g_2 \circ H$ and $q_2 \in g_1 \circ H$, (indeed, H is a subgroup so $h_2 \circ h_1^{-1}, h_1 \circ h_2^{-1} \in H$).

Now, take an arbitrary element $x \in g_1 \circ H$. Then $x = g_1 \circ h$ for some $h \in H$. Substituting form 3.5 we get

$$x = (g_2 \circ (h_2 \circ h_1^{-1})) \circ h = g_2 \circ (h_2 \circ h_1^{-1} \circ h) \text{ and so } x \in g_2 \circ H.$$

Indeed, H is a subgroup so $h_2 \circ h_1 \circ h$ belongs to H.

Similarly, one gets that any $z \in g_2 \circ H$ belongs to $g_1 \circ H$. So, we have shown that $g_1 \circ H = g_2 \circ H$.

H is a subgroup, so $e \in H$, and therefore $g \in g \circ H$ for every $g \in G$. This means that every element from G belongs to some $g' \circ H$. Hence, the system $\{g \circ H \,|\, g \in G\}$ forms a partition of G.

To finish the argument, we show that all sets $g \circ H$ have the same number of elements which is $k = |H|$. Denote $H = \{h_1, \ldots, h_k\}$. Then

$$g \circ H = \{g \circ h_1, \ldots, g \circ h_k\}.$$

If $g \circ h_i = g \circ h_j$ then $(g^{-1} \circ g) \circ h_i = (g^{-1} \circ g) \circ h_2$, which means that $h_i = h_j$ (see also 3.1.10).

We have shown that the set of n elements is divided into disjoint parts each of them having k elements. Hence n is divisible by k. (Note that there are n/k distinct sets $g \circ H$.) □

3.3.2 Order of a Finite Group. The number of elements of a finite group $(G, \circ, e)$ is often called its *order*. The above theorem can be formulated as follows: The order of any subgroup $(H, \circ, e)$ of a finite group $(G, \circ, e)$ divides the order of $(G, \circ, e)$.