Math 331 Fall 2026, Lecture 2 and 3: Defining semigroups, monoids and groups

website: `https://www.cip.ifi.lmu.de/~grinberg/t/26fa/`
(GPT-6 was used to polish the writing that follows.)

1. Groups

1.1. Definition

Definition 1.1.1. Let S be a set.

(a) A **binary operation** on S means a function f from $S \times S$ to S. That is, it means a function that takes two elements of S as inputs and returns an element of S. We will usually write this function in **infix notation**: i.e., instead of writing $f(s,t)$, we shall write $s \ f \ t$. We will often use symbols like $\cdot$, $+$ and $*$ instead of f, so our expressions will look like $s \cdot t$, $s + t$ or $s * t$.

For example, $+$ is a binary operation on the set $\mathbb{Z}$.

(b) A binary operation $*$ on S is said to be **associative** if all $a, b, c \in S$ satisfy

$$(a * b) * c = a * (b * c).$$

(c) If $*$ is an associative binary operation on S, then the pair $(S, *)$ is called a **semigroup**. (Alternatively, we can say that "S is a semigroup under $*$", or just "S is a semigroup" if the operation $*$ is inferrable from the context; but the operation $*$ is still part of the structure even if we don't explicitly name it.)

(d) Given a binary operation $*$ on S, a **neutral element** (also called an **identity element**) of $*$ means an element $e \in S$ such that every $a \in S$ satisfies

$$a * e = e * a = a.$$

We also call e a neutral element of S with respect to $*$, or just a neutral element of the pair $(S, *)$.

(e) A **monoid** means a semigroup $(S, *)$ that has a neutral element (i.e., for which there exists a neutral element $e \in S$ of $*$). We will soon show (Proposition 1.3.1) that this neutral element is unique.

(f) Now let $(S, *)$ be a monoid with neutral element e. An **inverse** of an element $a \in S$ (with respect to $*$) means an element $b \in S$ such that

$$a * b = b * a = e.$$

(g) A **group** means a monoid $(S, *)$ such that each element $a \in S$ has an inverse with respect to $*$. We will soon show (Proposition 1.3.3) that this inverse of a is unique.

Remark 1.1.2. We will mostly deal with groups in this course. Monoids are no less natural, but have fewer general properties (since there is "less to work with" without inverses). Semigroups occur even more frequently, but many of the useful examples we will encounter are already monoids.

Let us summarize the above definition as a "flowchart" for checking whether a given pair $(S, *)$ is a semigroup, monoid or group:

Given a pair $(S, *)$:

- **If** $*$ is a map from $S \times S$ to S, then $*$ is a binary operation on S.
  - •• **If, in addition**, $*$ is associative (that is, $(a * b) * c = a * (b * c)$ for all $a, b, c \in S$), then $(S, *)$ is a semigroup.
    - ••• **If, in addition,** $*$ has a neutral element e (that is, an element $e \in S$ such that all $a \in S$ satisfy $a * e = e * a = a$), then $(S, *)$ is a monoid.
      - •••• **If, in addition,** each $a \in S$ has an inverse (that is, some $b \in S$ such that $a * b = b * a = e$), then $(S, *)$ is a group.

1.2. Examples

Let us scour the mathematics we know for examples of semigroups, monoids and groups. We begin with the number systems from Lecture 1:

- The pair $(\mathbb{Z}_{>0}, +)$ (that is, the set $\mathbb{Z}_{>0}$ equipped with the binary operation $+$) is a semigroup, since $+$ is associative. It is not a monoid: the only possible neutral element would be 0, but $0 \notin \mathbb{Z}_{>0}$. Thus, it is not a group either.

  The same applies to $(\mathbb{Q}_{>0}, +)$ and $(\mathbb{R}_{>0}, +)$.

- The pair $(\mathbb{Z}, +)$ is a group: addition is associative, 0 is a neutral element, and each $a \in \mathbb{Z}$ has the inverse $-a \in \mathbb{Z}$.

  The same applies to $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ and $(\mathbb{C}, +)$.

- The pair $(\mathbb{Z}_{>0}, \cdot)$ (that is, the set $\mathbb{Z}_{>0}$ equipped with multiplication) is a monoid, since multiplication is associative and 1 is a neutral element. It is not a group, since 2 has no inverse in $\mathbb{Z}_{>0}$.

  Similarly, the pairs $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{R}, \cdot)$ and $(\mathbb{C}, \cdot)$ are also monoids but not groups. In each case, 0 has no inverse.

However, $(\mathbb{Q} \setminus \{0\}, \cdot)$ is a group, since products of nonzero rational numbers are nonzero and each nonzero rational number has a nonzero reciprocal. This group is denoted by $\mathbb{Q}^\times$. Similarly, the nonzero real numbers and the nonzero complex numbers form groups under multiplication, denoted by $\mathbb{R}^\times$ and $\mathbb{C}^\times$, respectively.

The Hamiltonian quaternions $\mathbb{H}$ give a similar example: nonzero quaternions have nonzero products and have inverses. Thus, $(\mathbb{H} \setminus \{0\}, \cdot)$ is a group, even though quaternion multiplication is not commutative.

- The pair $(\mathbb{Z}, -)$ is not a semigroup, since subtraction is not associative: the equality $(a-b)-c = a-(b-c)$ fails for any three integers a, b, c with $c \neq 0$.

  The pair $(\mathbb{Q}, /)$ is not a semigroup either. In fact, division is not even a binary operation on $\mathbb{Q}$, because division by 0 is undefined. Restricting to $\mathbb{Q} \setminus \{0\}$ does give a binary operation, but it is still not associative. For example, $(1/2)/3 = 1/6 \neq 3/2 = 1/(2/3)$.

  What about exponentiation on $\mathbb{Z}_{>0}$, viewed as the binary operation that sends (a, b) to a^b? This does not give a semigroup either, since associativity fails. For example, $2^{\left(2^3\right)} = 256 \neq 64 = \left(2^2\right)^3$.

- What about the pair $(\mathbb{Z}_{<0}, +)$, where $\mathbb{Z}_{<0} = \{-1, -2, -3, \ldots\}$ is the set of all negative integers? This is again a semigroup but not a monoid, since addition is associative but its only possible neutral element, 0, does not belong to $\mathbb{Z}_{<0}$.

  What about $(\mathbb{Z}_{<0}, \cdot)$? Here, multiplication is not even a binary operation on $\mathbb{Z}_{<0}$: products of negative integers are positive, so the output does not belong to the specified set. Thus, this pair is certainly not a semigroup, a monoid or a group.

- For any positive integer n, consider the set $\mathbb{R}^{n \times n}$ of all $n \times n$-matrices with real entries. It forms a group $(\mathbb{R}^{n \times n}, +)$ under addition. Its neutral element is the zero matrix, and the inverse of a matrix A is $-A$.

  Consider furthermore the set $\mathrm{GL}_n(\mathbb{R})$ of all invertible (= nonsingular) $n \times n$-matrices. This forms a group $(\mathrm{GL}_n(\mathbb{R}), \cdot)$ under multiplication. Its neutral element is I_n (the identity matrix), and its inverses are the usual matrix inverses. The product of two invertible matrices is again invertible, so multiplication is indeed a binary operation on this set.

  The pair $(\mathbb{R}^{n \times n}, \cdot)$ is a monoid but not a group: the identity matrix is still neutral, but not every matrix is invertible. For example, the zero matrix has no inverse. On the other hand, $(\mathrm{GL}_n(\mathbb{R}), +)$ is not even a semigroup, since sums of invertible matrices don't have to be invertible.

For two distinct positive integers n and m, two $n \times m$-matrices cannot be multiplied using ordinary matrix multiplication. Thus, there is no semigroup $(\mathbb{R}^{n \times m}, \cdot)$ under matrix multiplication.

Let us now move on to less familiar operations and less familiar sets.

- For any set A, consider the set $\operatorname{Map}(A, A) = A^A$ of all maps from A to A.

  For any two maps $f, g \in \operatorname{Map}(A, A)$, the composition $f \circ g$ also belongs to $\operatorname{Map}(A, A)$. Recall that $f \circ g$ sends an element a to $f(g(a))$: we apply g first and then f. Thus, the composition operation $\circ$ is a binary operation on $\operatorname{Map}(A, A)$. It is associative, since both $(f \circ g) \circ h$ and $f \circ (g \circ h)$ send each $a \in A$ to $f(g(h(a)))$.

  Therefore, $(\operatorname{Map}(A, A), \circ)$ is a semigroup. The identity map $\mathrm{id}_A : A \to A$ (which sends each $a \in A$ to a itself) is a neutral element, so this semigroup is a monoid.

  Is it a group? This is asking whether every map $f : A \to A$ is invertible. This is true when A is empty or a 1-element set (because in either case, the only map from A to A is its identity map). If A has at least two elements, however, a constant map from A to A is not invertible. Thus, in all these other cases, $(\operatorname{Map}(A, A), \circ)$ is not a group.

  To "fix" this, we can restrict ourselves to the bijective (equivalently, invertible) maps from A to A. These are called the **permutations** of A, and they do form a group under composition: compositions and inverses of bijections are again bijections, and the identity map is a bijection. This group is called the **symmetric group** of A, and we will spend a while studying its properties.

- Fix a set A. Consider the power set of A; this is the set $\mathcal{P}(A)$ of all subsets of A. For example, if $A = \{1, 2, 3\}$, then

$$\mathcal{P}(A) = \{\varnothing,\ \{1\},\ \{2\},\ \{3\},\ \{1,2\},\ \{2,3\},\ \{1,3\},\ \{1,2,3\}\}.$$

  We would like to make $\mathcal{P}(A)$ into a semigroup, or better yet a monoid, or even a group. What operations can we use?

  Two well-known operations are $\cup$ and $\cap$.

  The pair $(\mathcal{P}(A), \cup)$ is a semigroup, since $(X \cup Y) \cup Z = X \cup (Y \cup Z)$ for all $X, Y, Z \subseteq A$. It is also a monoid, with neutral element $\varnothing$. Unless $A = \varnothing$, it is not a group: a nonempty subset X has no inverse, since $X \cup Y$ can never be empty.

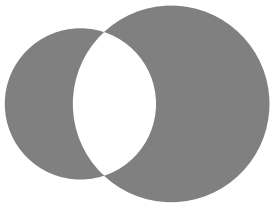
  The pair $(\mathcal{P}(A), \cap)$ is also a semigroup, since $(X \cap Y) \cap Z = X \cap (Y \cap Z)$ for all $X, Y, Z \subseteq A$. It is a monoid with neutral element A. Unless $A = \varnothing$,

it is not a group: the empty set has no inverse, since $\varnothing \cap Y = \varnothing \neq A$ for every $Y \subseteq A$. When $A = \varnothing$, both of these monoids are one-element groups.

Can you think of a binary operation on $\mathcal{P}(A)$ that does make it into a group for every A? Mason suggests the operation $\triangle$ defined by

$$X \triangle Y = (X \cup Y) \setminus (X \cap Y) = (X \setminus Y) \cup (Y \setminus X).$$

This is called the **symmetric difference** of X and Y. Here is the Venn diagram of $X \triangle Y$:

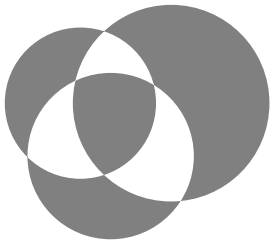


(where the two circles represent X and Y, while the shaded region is $X \triangle Y$).

Does this operation give a semigroup? Yes, if we can show that $\triangle$ is associative. This can be done using truth tables: for any $X, Y, Z \subseteq A$, we must show that a given element $p \in A$ lies in $(X \triangle Y) \triangle Z$ if and only if it lies in $X \triangle (Y \triangle Z)$. There are 8 cases to consider:

- the case $p \in X$, $p \in Y$ and $p \in Z$;
- the case $p \in X$, $p \in Y$ and $p \notin Z$;
- and so on (6 more cases).

You can visualize this proof through a Venn diagram:



(where the three circles represent X, Y and Z, while the shaded region is $(X \triangle Y) \triangle Z = X \triangle (Y \triangle Z)$). A nicer proof observes that both $(X \triangle Y) \triangle Z$ and $X \triangle (Y \triangle Z)$ consist of precisely those $p \in A$ that belong to an **odd** number of the sets X, Y, Z (that is, either to exactly one or to all three).

Thus, $(\mathcal{P}(A), \triangle)$ is a semigroup. It is a monoid with neutral element $\varnothing$. Is it a group? Yes: the inverse of any $X \in \mathcal{P}(A)$ is X itself, since

$$X \triangle X = (X \cup X) \setminus (X \cap X) = X \setminus X = \varnothing.$$

- Consider the set of all Euclidean transformations of the plane (also called rigid transformations of the plane). These are the invertible maps from the plane $\mathbb{R}^2$ to itself that preserve distances. Examples are parallel shifts (translations), rotations and reflections, as well as their compositions. They form a group under composition: the identity map preserves distances, and so do compositions and inverses of distance-preserving maps.

  What happens if we restrict to Euclidean transformations that preserve orientation? These still form a group under composition.

  What about the Euclidean transformations that reverse orientation? They do not even form a semigroup under composition, since composing two such transformations gives an orientation-preserving transformation, not an orientation-reversing one!

- Consider a weird-looking binary operation $*$ on the set $\mathbb{Q}$ defined by

$$a * b = ab + a + b \qquad \text{for all } a, b \in \mathbb{Q}.$$

  Is $(\mathbb{Q}, *)$ a semigroup? Is it a monoid? Is it a group?

  – It is a semigroup. Indeed, associativity can be checked directly: for all $a, b, c \in \mathbb{Q}$, we have

$$\begin{aligned}(a * b) * c &= (ab + a + b) * c \\ &= (ab + a + b)\, c + (ab + a + b) + c \\ &= abc + ab + ac + bc + a + b + c,\end{aligned}$$

  and expanding $a * (b * c)$ gives the same result.

  – It is a monoid. The element 0 is neutral, since $0 * a = 0a + 0 + a = a$ and likewise $a * 0 = a$ for every $a \in \mathbb{Q}$.

  – It is not a group. The element -1 has no inverse, because $(-1) * a = -a - 1 + a = -1$ for every $a \in \mathbb{Q}$, whereas an inverse would have to satisfy $(-1) * a = 0$.

  All other rational numbers do have inverses: for $a \neq -1$, the inverse is $\dfrac{-a}{a+1}$. This can be checked directly from the definition of $*$.

  Thus, $(\mathbb{Q}, *)$ is a monoid. But actually, it is just the monoid $(\mathbb{Q}, \cdot)$ with its elements renamed – namely, with each rational number c renamed as $c - 1$. Indeed, for three rational numbers a, b, c, we have

$$c - 1 = (a - 1) * (b - 1) \qquad \text{if and only if} \qquad c = ab.$$

  Equivalently,

$$c = a * b \qquad \text{if and only if} \qquad c + 1 = (a + 1)\,(b + 1)\,.$$

So $(\mathbb{Q}, *)$ is a monoid but not really a new one: it is just a version of $(\mathbb{Q}, \cdot)$ with its elements renamed. We say that it is **isomorphic** to $(\mathbb{Q}, \cdot)$. We will give a formal definition of isomorphism later.

- Consider the set $\{0\}$ with the binary operation $*$ defined by $0 * 0 = 0$. The pair $(\{0\}, *)$ is a group: its neutral element is 0, which is also its own inverse. More generally, any one-element set becomes a group when we define its binary operation in the only possible way. Such a group is called a **trivial group**.

- What would a group $(G, *)$ with two elements look like? It must have a neutral element – let's call it e. Let f be the other element of G. Since e is neutral, we have $e * e = e$, $e * f = f$ and $f * e = f$.

  What can $f * f$ be? The only possibilities are e and f. But f must have an inverse, and this inverse cannot be e, since $f * e = f \neq e$. Therefore, the inverse of f is f itself, and hence $f * f = e$. This uniquely determines the multiplication table (more precisely: the $*$-table) of G:

$x * y$	$y = e$	$y = f$
$x = e$	e	f
$x = f$	f	e

  .

  Conversely, if we start with a two-element set $\{e, f\}$ and define a binary operation $*$ by this table, then we do get a group. We can verify the axioms directly (start by showing that e is a neutral element), but a shortcut makes the associativity check especially quick: Whenever a binary operation has a neutral element e, the equality $(x * y) * z = x * (y * z)$ holds immediately if **any** of x, y, z equals e. Thus, in our case, the only remaining check is

$$(f * f) * f = e * f = f = f * e = f * (f * f).$$

  The table also shows that e is neutral and that both elements are their own inverses, so all the group axioms hold.

  Alternatively (thanks to Nico Ghiron), we can recall that the two numbers 1 and -1 multiply by the same rule as the elements e and f above:

xy	$y = 1$	$y = -1$
$x = 1$	1	-1
$x = -1$	-1	1

  .

  Thus, we already have a group $(\{1, -1\}, \cdot)$ with two elements! The group $(\{e, f\}, *)$ we are constructing is just an isomorphic copy of this group: it is the same group with 1 and -1 renamed as e and f, and with the operation $\cdot$ renamed as $*$.

1.3. Some first general properties of semigroups, monoids and groups

So we have defined three rather general classes of objects (semigroups, monoids and groups) and given some examples and non-examples for each. Let us now prove a few things about these objects in general, just using their definitions.

First of all, neutral elements – when they exist – are unique:

Proposition 1.3.1. Let $*$ be a binary operation on a set S. Then, $*$ has **at most one** neutral element. In other words, if a neutral element exists, then it is unique.

Proof. We must show that any two neutral elements of $*$ are equal.

Let e and f be two neutral elements of $*$. We want to prove that $e = f$. Luke suggests the following argument: Since f is neutral, $e * f = e$; since e is neutral, $e * f = f$. Comparing these equalities, we get $e = f$.

Thus, any two neutral elements of $*$ are equal, so there is at most one. □

Definition 1.3.2. Let $(S, *)$ be a monoid. Proposition 1.3.1 allows us to speak of **the neutral element** of $(S, *)$ without having to worry about potentially confusing several different neutral elements. We shall do so, and we will usually denote this neutral element by e.

In some of the later lectures, we will rename the operation $*$ as $\cdot$ (and simply write st for $s * t$), and correspondingly rename the neutral element e as 1, thus pretending that the operation $*$ is some kind of multiplication. Of course, this notation should not be used when the set S already has a different multiplication defined on it (or contains the number 1 but this number is not its neutral element); but it works well when we are dealing with a generic monoid $(S, *)$ that we know nothing about.

The same applies to inverses in a monoid:

Proposition 1.3.3. Let $(S, *)$ be a monoid, and let $a \in S$ be an element that has an inverse. Then, this inverse is unique.

Proof. We must show that any two inverses of a are equal. Let x and y be two such inverses. Then, $x * a = e$ and $a * y = e$ (where e is the neutral element of $(S, *)$). But associativity yields $(x * a) * y = x * (a * y)$. In view of $x * \underbrace{(a * y)}_{=e} = x * e = x$ (since e is neutral) and $\underbrace{(x * a)}_{=e} * y = e * y = y$ (again since e is neutral), we can rewrite this as $x = y$. Thus, we have shown that any two inverses of a are equal. Since an inverse exists by assumption, it is unique. □

Definition 1.3.4. Let $(S, *)$ be a monoid. Let $a \in S$. Proposition 1.3.3 tells us that if a has an inverse, then we can call it **the inverse** of a without worrying that this might be ambiguous. We shall do so, and we will denote this inverse by a^{-1}.

The following properties of inverses are crucial for any work with them. You might have already encountered them when dealing with inverses of matrices and maps; here we state them in an arbitrary monoid (thus generalizing the matrix case and partly generalizing the case of maps[1]).

Theorem 1.3.5 (rules for inverses)**.** Let $(S, *)$ be a monoid with neutral element e.

(a) If $a \in S$ has an inverse a^{-1}, then this inverse a^{-1} itself has an inverse, namely

$$\left(a^{-1}\right)^{-1} = a.$$

(b) If $a, b \in S$ both have inverses, then $a * b$ also has an inverse, namely

$$(a * b)^{-1} = b^{-1} * a^{-1}.$$

This is known as the **socks-and-shoes rule**. It is analogous to the rule $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$ for invertible maps and the rule $(AB)^{-1} = B^{-1}A^{-1}$ for invertible square matrices of the same size.

Proof.

(a) Since a^{-1} is an inverse of a, we have $a * a^{-1} = a^{-1} * a = e$. The same equalities show that a is an inverse of a^{-1}. Thus, $\left(a^{-1}\right)^{-1} = a$.

(b) We have to show that $b^{-1} * a^{-1}$ is an inverse of $a * b$. For this, we must prove that

$$\left(b^{-1} * a^{-1}\right) * (a * b) = e \qquad \text{and}$$
$$(a * b) * \left(b^{-1} * a^{-1}\right) = e.$$

We check both of these equalities now. Using associativity in the first two

[1]Only partly, because maps between different sets do not live in a monoid.

steps of each calculation, we obtain

$$\begin{aligned}\left(b^{-1} * a^{-1}\right) * (a * b) &= b^{-1} * \left(a^{-1} * (a * b)\right) \\ &= b^{-1} * \left(\underbrace{\left(a^{-1} * a\right)}_{=e} * b\right) \\ &= b^{-1} * \underbrace{(e * b)}_{=b} = b^{-1} * b = e\end{aligned}$$

and

$$\begin{aligned}(a * b) * \left(b^{-1} * a^{-1}\right) &= a * \left(b * \left(b^{-1} * a^{-1}\right)\right) \\ &= a * \left(\underbrace{\left(b * b^{-1}\right)}_{=e} * a^{-1}\right) \\ &= a * \underbrace{\left(e * a^{-1}\right)}_{=a^{-1}} = a * a^{-1} = e.\end{aligned}$$

Thus, $b^{-1} * a^{-1}$ is the inverse of $a * b$. □

In this proof, we had to deal with the annoyance of regrouping products using associativity. It would have been much nicer if we had been able to write products like

$$a * b * b^{-1} * a^{-1}$$

without parentheses. But $*$ is a binary operation: it has only two inputs. Thus, it is not a-priori clear that we can make sense of "products" of several elements, such as $a * b * c * d * e$, without parentheses.

For example, the "product" $a * b * c * d$ in a semigroup can be fully parenthesized in 5 ways:

$$\begin{gathered}((a * b) * c) * d, \\ (a * (b * c)) * d, \\ a * (b * (c * d)), \\ a * ((b * c) * d), \\ (a * b) * (c * d).\end{gathered}$$

Do these 5 ways all produce the same result?

For a product $a * b * c$, there are only 2 ways to insert parentheses: $(a * b) * c$ and $a * (b * c)$. They produce the same result by associativity.

For $a * b * c * d$, the answer is not quite as immediate. Nevertheless, repeated use of associativity shows that all 5 ways give the same result:

$$
\begin{aligned}
&(a * (b * c)) * d \\
&= ((a * b) * c) * d && \text{(by associativity for } a, b \text{ and } c) \\
&= (a * b) * (c * d) && \text{(by associativity for } a * b, c \text{ and } d) \\
&= a * (b * (c * d)) && \text{(by associativity for } a, b \text{ and } c * d) \\
&= a * ((b * c) * d) && \text{(by associativity for } b, c \text{ and } d)\,.
\end{aligned}
$$

Does this generalize to larger products? Next time we will see the answer.