# Perspectives on Abstract Algebra

Sebastian Ørsted

November 26, 2016

In this note, I expand upon the coverage of a selected number of concepts from Lauritzen (2003). The main purpose is to provide additional examples and explanations, hopefully providing the reader with more intuition about the abstract constructions we deal with in the subject. I plan to expand the notes throughout the course as I notice more aspects of the theory I wish to emphasize further. The newest version will always be available at

`http://home.math.au.dk/sorsted/persp_algebra.pdf`

I shall allow myself to use some notation which is universal standard: I write $\varphi\colon X \hookrightarrow Y$ to indicate that the map $\varphi$ is injective. The hooked arrow is meant to suggest an arrow put together with a subset symbol "$\subset$". This makes sense because an injective map can be thought of as a kind of inclusion; indeed, notice, for instance, that for an injective group homomorphism $i\colon G \hookrightarrow H$, the image $\operatorname{Im}(i) \subset H$ is a subgroup isomorphic to $G$ via $i$. Because of this, injective homomorphisms are often called **embeddings**. Similarly, we write $\varphi\colon X \twoheadrightarrow Y$ to indicate that the map $\varphi$ is surjective. My best guess is that this notation is meant to suggest "hit $Y$ extra hard".

# Chapter 1

# Groups

## 1.1 Quotient groups

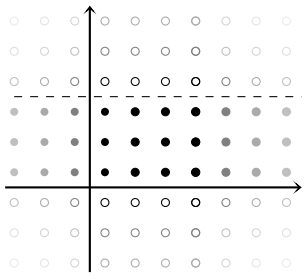
Quotient groups take time getting used to, since they provide perhaps one of the most abstract concepts in elementary group theory. However, the power they represents can hardly be overestimated; indeed, they are an invaluable tool for creating new groups with all kinds of interesting properties. We shall see later on how quotient groups allow us to give *presentations* of groups, which provides a technique to describe *any* group by means of a collection of *generators* and a set of *relations* to impose on these generators.

Given a normal subgroup $N$ of a group $G$, the quotient group comes with a surjective group homomorphism $G \twoheadrightarrow G/N$, $g \mapsto [g] = gN$, called the **canonical map** or the **quotient map**, which we shall usually denote $\pi$. When calculating in the quotient group, we have $[g][h] = [gh]$, so that the group operation more or less reflects the group operation on $G$. However, there is the important difference that $[n] = [e] = e \in G/N$ for all $n \in N$, showing that all elements in $N$ become the trivial element (that is, $e$) in the quotient group $G/N$. In other words, we have the following interpretation of quotient groups:

**Interpretation 1.1.** *The quotient group $G/N$ is the group $G$ where all elements of $N$ are declared to be trivial (i.e. equal to the neutral element $e$). In other words, when calculating in $G/N$, we calculate as in $G$, but ignore elements of $N$.*

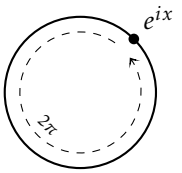
**Example 1.2.** The groups $\mathbb{Z}/d\mathbb{Z}$ is the quotient group of $\mathbb{Z}$ by the normal subgroup $d\mathbb{Z} \subset \mathbb{Z}$. In other words, it is the group $\mathbb{Z}$ where we simply ignore all multiples of $d$. ○

**Example 1.3.** Suppose we would like to create a group which can be visualized as a grid like on the right. We would like it to be similar to the group $\mathbb{Z}^2$, with addition as group operation, but with one important difference: Every point is identified with the point three steps above it. In other words, the same three rows of points repeat themselves indefinitely and are equal to the three rows between the first axis and the dashed line. We can describe this group as the



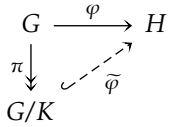
group $\mathbb{Z}^2$ where we ignore multiples of the group element $(0,3) \in \mathbb{Z}^2$. In other words, the group we want is nothing but the quotient group of $\mathbb{Z}^2$ by the subgroup $\mathbb{Z} \cdot (0,3) \subset \mathbb{Z}^2$ (notice that we did not have to check that this subgroup was normal since $\mathbb{Z}^2$ is Abelian, see Lauritzen 2003, Exercise 2.14).○

**Example 1.4.** Let $S^1 \subset \mathbb{C}$ be the unit circle in the complex plane. Each point in $S^1$ has the form $e^{ix}$ for some real number $x$, showing that the circle is a group under multiplication, $e^{ix}e^{iy} = e^{i(x+y)}$. The choice of $x$ in the formula $e^{ix}$ is not unique; indeed, adding any multiple of $2\pi$ to $x$ produces the same point on the unit circle. In other words, we can calculate mod $2\pi$ on $S^1$, ignoring any multiple of this number. This is the intuitive reason for the group isomorphism $S^1 \cong \mathbb{R}/2\pi\mathbb{Z}$. A formal argument using the Isomorphism Theorem may be found in Example 2.5.2, Lauritzen (ibid.). ○




Let us restate the Isomorphism Theorem in a slightly different and, in some cases, more useful form, using the notation mentioned in the introduction. We also prefer to call it Noether's First Isomorphism Theorem, after the German mathematician Emmy Noether (1882–1935).

**Noether's First Isomorphism Theorem 1.5.**
*If $\varphi\colon G \to H$ is a group homomorphism with kernel $K = \mathrm{Ker}(\varphi)$, then there exists a well-defined map $\widetilde{\varphi}\colon G/K \hookrightarrow H$ given by $\widetilde{\varphi}([g]) = \varphi(g)$ which is an injective group homomorphism. In terms of the canonical map $\pi\colon G \twoheadrightarrow G/K$, this means that $\varphi = \widetilde{\varphi} \circ \pi$, so that the diagram on the right is commutative.*




The proof is the same as in the course textbook, and we omit it.

**Example 1.6.** The differential map can be regarded as a group homomorphism $C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ on the additive Abelian group of all smooth (i.e., infinitely often differentiable) functions on $\mathbb{R}$. Its kernel is the subgroup $\mathbb{R} \subset C^\infty(\mathbb{R})$ of constant functions, which is normal because $C^\infty(\mathbb{R})$ is Abelian (cf. Exercise 2.14 in ibid.). Thus the Isomorphism Theorem provides us with an injective map $\tilde{d}\colon C^\infty(\mathbb{R})/\mathbb{R} \hookrightarrow C^\infty(\mathbb{R})$ given by $[f] \mapsto df$. ○

**Advanced Example 1.7.** Let us like in Example 1.6 consider the additive Abelian group $C^\infty(\mathbb{R})$ of smooth functions on $\mathbb{R}$ and $d$ the differential operator. If $\mathbb{R}[x] \subset C^\infty(\mathbb{R})$ denotes the subgroup of all polynomial functions, this group is again normal because $C^\infty(\mathbb{R})$ is Abelian. Consider the composition

$$C^\infty(\mathbb{R}) \xrightarrow{d} C^\infty(\mathbb{R}) \twoheadrightarrow C^\infty(\mathbb{R})/\mathbb{R}[x]$$

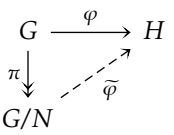
of $d$ with the canonical map $C^\infty(\mathbb{R}) \twoheadrightarrow C^\infty(\mathbb{R})/\mathbb{R}[x]$. Its kernel consists exactly of the functions $f \in C^\infty(\mathbb{R})$ such that $df \in \mathbb{R}[x]$. But a function whose derivative is a polynomial is itself a polynomial, hence the kernel is $\mathbb{R}[x]$. Applying the First Isomorphism Theorem we obtain an injective map $C^\infty(\mathbb{R})/\mathbb{R}[x] \to C^\infty(\mathbb{R})/\mathbb{R}[x]$ which we shall denote $D$, given by $D[f] = [df]$. We claim that $D$ is in fact an isomorphism of groups (the reader can then check that it is even an isomorphism of vector spaces), so let us prove surjectivity. For this we simply

note that any smooth (or even continuous) function has an antiderivative according to the Fundamental Theorem of Calculus. Thus given $[f] \in C^\infty(\mathbb{R})$ we can always find $[g] \in C^\infty(\mathbb{R})$ with $D[g] = [dg] = [f]$. This proves that $D$ is indeed an isomorphism.

This allows us to define an inverse $\int : C^\infty(\mathbb{R})/\mathbb{R}[x] \to C^\infty(\mathbb{R})/\mathbb{R}[x]$, mapping a class $[f]$ to the class $[g]$ of some antiderivative $g$ of $f$. By abstract algebraic nonsense, we have removed the obstacle that the antiderivative of a function is not unique. It comes at a price, though; as the reader may verify on concrete examples, multiplication is *not* a well-defined operation on $C^\infty(\mathbb{R})/\mathbb{R}[x]$. It *does* make sense to multiply by *polynomials* or even compose on the right with polynomials, but that is essentially all you can do. ○

Actually, there exists a stronger version of the First Isomorphism Theorem, allowing us to take quotients by any normal subgroup contained in the kernel.

**Fundamental Homomorphism Theorem 1.8.**
*If $\varphi : G \to H$ is a group homomorphism and $N$ a normal subgroup of $G$ contained in* $\operatorname{Ker}(\varphi)$*, then then exists a well-defined map $\widetilde{\varphi} : G/N \to H$ given by $\widetilde{\varphi}([g]) = \varphi(g)$ which is a group homomorphism with kernel* $\operatorname{Ker}(\varphi)/N$*. In terms of the canonical map $\pi : G \twoheadrightarrow G/N$, this means that $\varphi = \widetilde{\varphi} \circ \pi$, so that the diagram on the right is commutative.*




Notice that since $N$ is a normal subgroup of $G$ contained in $\operatorname{Ker}(\varphi)$, it is a normal subgroup of $\operatorname{Ker}(\varphi)$ (why?), so the quotient group $\operatorname{Ker}(\varphi)/N$ makes sense. Note also that the assumption that $N \subset \operatorname{Ker}(\varphi)$ is equivalent to requiring that $\varphi$ vanishes on $N$, that is, $\varphi(n) = e$ for all $n \in N$.

***Proof.*** To prove that the map $\widetilde{\varphi}$ is well-defined, suppose $gN = hN$ in $G/N$ and let us prove that $\varphi(g) = \varphi(h)$. From Lemma 2.2.6(ii) in the course textbook, we have $gh^{-1} \in N \subset \operatorname{Ker}(\varphi)$, hence

$$e = \varphi(gh^{-1}) = \varphi(g)\varphi(h)^{-1},$$

showing that indeed $\varphi(g) = \varphi(h)$. To determine the kernel of $\widetilde{\varphi}$, suppose $\widetilde{\varphi}(gN) = \varphi(g) = e$. Then we have $g \in \operatorname{Ker}(\varphi)$, and thus $gN \in \operatorname{Ker}(\varphi)/N$. The proof that $\widetilde{\varphi}$ is a group homomorphism is left to the reader. □

**Advanced Example 1.9.** If $G$ is a group, there is a way to "make $G$ Abelian" while maintaining as much of the structure of $G$ as possible. Namely, there exists a group $G^{\mathrm{Ab}}$, the **Abelianization** of $G$, which can be thought of abstractly as the group $G$ where we simply add the relation $gh = hg$ to the existing relations of $G$.

To define $G^{\mathrm{Ab}}$ formally, consider the **commutator**

$$[g,h] = (gh)(hg)^{-1} = ghg^{-1}h^{-1}, \qquad g,h \in G.$$

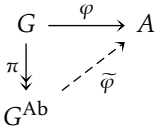
Notice that $G$ is Abelian if and only if $[g,h] = e$ for all $g,h \in G$. We then consider the smallest normal subgroup containing all commutators. Formally, if $R = \{[g,h] \mid g,h \in G\}$ is the set of all commutators, we define

$$N(R) = \bigcap_{N \supset R} N, \tag{1.1}$$

where the intersection runs over all normal subgroups $N$ of $G$ containing $R$. As the reader may check, the intersection of any number of normal subgroups is itself a normal subgroup. Now we simply define $G^{\mathrm{Ab}} = G/N(R)$ and obtain that

$$[e] = [[g,h]] = [g][h][g]^{-1}[h]^{-1} = ([g][h])([h][g])^{-1}$$

for all $[g],[h] \in G^{\mathrm{Ab}}$. It follows that $[g][h] = [h][g]$, proving that $G^{\mathrm{Ab}}$ is Abelian. If the group $G$ was Abelian to begin with, all commutators were already trivial, and we have $G^{\mathrm{Ab}} = G$.




The Abelianization $G^{\mathrm{Ab}}$ comes with a surjective homomorphism $\pi\colon G \twoheadrightarrow G^{\mathrm{Ab}}$, namely the canonical map. This has the following universal property: If $\varphi\colon G \to A$ is any homomorphism from $G$ to an Abelian group $A$, there exists a map $\widetilde{\varphi}\colon G^{\mathrm{Ab}} \to A$ satisfying $\varphi = \widetilde{\varphi} \circ \pi$, that is, the diagram on the right is commutative. This can be shown directly or follows from the Fundamental Homomorphism Theorem because the fact that $A$ is Abelian implies that $\varphi$ vanishes on all commutators and hence $R \subset \mathrm{Ker}(\varphi)$. Then (1.1) and the normality of $\mathrm{Ker}(\varphi)$ show that $N(R) \subset \mathrm{Ker}(\varphi)$, allowing us to apply the theorem with $N = N(R)$. We may interpret this by saying that $G^{\mathrm{Ab}}$ is, in a sense, the most general Abelian group with a map $G \to G^{\mathrm{Ab}}$. ○

**Remark 1.10.** Notice that the construction from (1.1) works for any subset $R$ of any group $G$. So we can always define the normal subgroup $N(R) \subset G$ generated by any collection of elements $R$. △

**Measure-Theoretic Example 1.11.** Given a measure space $(X, \mathscr{E}, \mu)$, the vector space $\mathscr{L}^1(\mu)$ of measurable functions $f\colon X \to \mathbb{C}$ satisfying $\int_X |f|\,d\mu < \infty$ has the defect that the integral cannot detect the difference between functions that agree almost everywhere. To make up for this defect, notice that the integral map $\int\colon \mathscr{L}^1(\mu) \to \mathbb{C}$ is linear and hence a group homomorphism, and that the kernel contains the (normal, since $\mathscr{L}^1(\mu)$ is Abelian) subgroup $N$ of all functions that are zero almost everywhere. Applying the Fundamental Homomorphism Theorem, we get an induced map $\mathscr{L}^1(\mu)/N \to \mathbb{C}$ which we also write as an intergral. We then arrive at the well-known space $L^1(\mu) = \mathscr{L}^1(\mu)/N$.

We have only argued that $L^1(\mu)$ is a group and the integral an additive group homomorphism. To show that $L^1(\mu)$ also has the structure of a *vector space* with the integral being a linear map, we simply note that there exists a similar Fundamental Homomorphism Theorem for vector spaces and linear maps, with the same proof. In fact, it is a special case of the module-theoretic version covered in the next chapter. ○

As the reader may have guessed, there are more Noether Isomorphism Theorems:

**Noether's Second Isomorphism Theorem 1.12.**
*Let $G$ be a group, $H$ a subgroup, and $N$ a normal subgroup. Then $HN$ is a subgroup of $G$, $H \cap N$ is a normal subgroup of $H$, and there exists an isomorphism of groups*

$$H/(H \cap N) \xrightarrow{\sim} (HN)/N, \qquad h(H \cap N) \longmapsto hN.$$

***Proof.*** The proof of the first two claims is left to the reader. For the isomorphism, we note that there exists a map $H \to (HN)/N$, mapping $h$ to $hN$. This map is surjective because any class $hnN \in (HN)/N$ satisfies $hnN = hN$. Hence we are done by the First Isomorphism Theorem if we can argue that the kernel is exactly $H \cap N$. But an $h \in H$ satisfies $hN = eN$ if and only if $h$ also belongs to $N$, which proves the claim. □

**Noether's Third Isomorphism Theorem 1.13.**
*Let $G$ be a group and $N$ a normal subgroup. Then we have:*

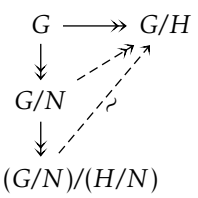
(i) *If $H$ is a subgroup of $G$ containing $N$, then $H/N$ is a subgroup of $G/N$. Furthermore, every subgroup of $G/N$ has this form.*

(ii) *If $H$ is a normal subgroup of $G$ containing $N$, then $H/N$ is a normal subgroup of $G/N$. Furthermore, every normal subgroup of $G/N$ has this form. Finally, we have an isomorphism*

$$(G/N)/(H/N) \xrightarrow{\sim} G/H, \qquad (gN)(H/N) \longmapsto gH.$$

The isomorphism is easy to remember because it resembles the cancellation rule in fractions. In fact, this may be the original motivation for the quotient notation.

***Proof.*** Most of the proof consists of a checking definitions and is covered in Exercise 2.30 in Lauritzen (2003). So we shall concentrate on the isomorphism. First note that $H \supset N$ implies that $N$ is contained in the kernel of the canonical map $G \twoheadrightarrow G/H$, so the Fundamental Homomorphism Theorem provides us with a map $G/N \twoheadrightarrow G/H$ with kernel $H/N$. Now applying the First Isomorphism Theorem to this map yields the desired isomorphism. □




## 1.2 Group actions

Recall the notion of a group action. Lauritzen (ibid.) defines an action of a group $G$ on a set $S$ to be a map

$$G \times S \longrightarrow S, \qquad (g, s) \longmapsto g \cdot s,$$

subject to the axioms

- $e \cdot s = s$ for all $s \in S$, and
- $(gh) \cdot s = g \cdot (h \cdot s)$ for all $g, h \in G$ and $s \in S$.

While it is certainly not evident at first, group actions are of fundamental importance in the theory of groups. Indeed, groups were introduced in the first place with the very purpose of defining group actions.

In this note, we present a different, but entirely equivalent definition of group actions. To motivate it, suppose we are given an action $G \times S \to S$ as defined above. Now we can associate to each $g \in G$ the map

$$\rho(g)\colon S \longrightarrow S, \qquad s \longmapsto g \cdot s. \tag{1.2}$$

This map is invertible with inverse $\rho(g^{-1})$, hence $\rho(g)$ belongs to $\mathrm{Bij}(S)$, the group of all bijective maps $S \to S$ (the group operation being composition of maps). Thus we have defined a map $\rho\colon G \to \mathrm{Bij}(S)$.

**Proposition 1.14.** *The map $\rho\colon G \to \mathrm{Bij}(S)$ defined in (1.2) is a group homomorphism.*

***Proof.*** If $g, h \in G$, then $\rho(gh)$ maps $s$ to $(gh)\cdot s$. But using the axioms of a group action, this is equal to $g\cdot(h\cdot s) = (\rho(g)\circ\rho(h))(s)$, proving that $\rho(gh) = \rho(g)\circ\rho(h)$, as claimed. □

Now we arrive at our equivalent definition: A **group action** of a group $G$ on a set $S$ is a group homomorphism $\rho\colon G \to \mathrm{Bij}(S)$. We usually simplify our notation and write $g\cdot s$, $gs$, or even $g(s)$ instead of $\rho(g)(s)$ when the map $\rho$ is clear from the context.

Many sources prefer to give this definition, and indeed, it is often easier to work with. This will become more clear when we discuss *representation theory* later in the course, which is, roughly speaking, the study of group actions on vector spaces. To go from Lauritzen's definition of group actions to ours, simply define $\rho$ by (1.2). To go from our definition to Lauritzen's, define the map $G\times S \to S$ by $(g,s)\mapsto \rho(g)(s)$.

**Example 1.15.** For any set $S$, there is an one important group acting on it, namely the group $G = \mathrm{Bij}(S)$: One simply lets $\rho\colon G \to \mathrm{Bij}(S)$ be the identity map. In a sense, this is the most general group action on $S$, since all other groups $G'$ act on $S$ by means of a map $\rho'\colon G' \to G$. Taking $S$ to be $G$ itself, we arrive at an action of $G$ on $G$ itself, namely $\rho(g)\colon G \to G$ given by $x \mapsto gx$. ○

**Example 1.16.** If $S$ is a finite set with $n$ elements, we may enumerate the elements by the numbers $1, 2, \ldots, n$. Hence we may as well assume $S = \{1, 2, \ldots, n\}$. Then the group $\mathrm{Bij}(S)$ of bijections on $S$ is simply the symmetric group $S_n$. In other words, an action of *any* group $G$ on $S$ is the same as a map $\rho\colon G \to S_n$.

In particular, applying Example 1.15, we arrive at the action of $S_n$ on the set $\{1, 2, \ldots, n\}$ described in Example 2.10.3(i) in Lauritzen (ibid.). ○

**Example 1.17.** The **trivial action** of a group $G$ on a set $S$ is the map $\rho\colon G \to \mathrm{Bij}(S)$ given by $\rho(g) = \mathrm{id}_S$ for all $g$, the identity map on $S$. In the simplified notation, this means that $gs = s$ for all $g \in G$. In other words, the action "does not do anything". ○

**Example 1.18.** If $V$ is a vector space over some field $k$, let $G = \mathrm{GL}(V)$ be the group of all bijective *linear* maps $V \to V$. This group acts in a natural way on $V$ by letting $\rho\colon G \to \mathrm{Bij}(S)$ be the inclusion. Thus for all $\varphi \in \mathrm{GL}(V)$, we have $\rho(\varphi)(v) = \varphi(v)$ for $v \in V$.

We could equivalent have chosen to work with matrices. Then we have the action of the group $\mathrm{GL}_n(k)$ of all invertible $n\times n$ matrices with entries in $k$ on the $k$-vector space $k^n$. ○

**Example 1.19.** In week 8, we considered the following exercise: Suppose $G$ is a group with $|G| > 2$ acting non-trivially (see Example 1.17) on a set $S$ with two elements. Prove that $G$ cannot be simple.

To do this, we may as well assume as in Example 1.16 that $S = \{1, 2\}$. Then as in that example, our action becomes a map $\rho\colon G \to S_2$. We claim that the normal subgroup $\mathrm{Ker}(\rho)$ is non-trivial (different from $\{e\}$ and $G$), which will prove that the group is not simple.

To verify our claim, note that since the action is non-trivial, this map $\rho$ cannot be constantly equal to the identity map, and since $S_2$ has two elements, this implies that $\rho$ is surjective. Thus $\mathrm{Ker}(\rho) \neq G$. To prove that $\mathrm{Ker}(\rho) \neq \{e\}$, suppose the opposite were true and apply Proposition 2.4.9 to obtain that $\rho$ is an injective map $\rho\colon G \to S_2$, contradicting that $|G| > 2 = |S_2|$. ○

Having seen these examples, we are one step closer to understanding what a group actually *is*. As we remarked briefly in the beginning, the historical motivation for the concept of groups was to make sense of group actions. As such, the idea of group actions is older than the idea of groups itself. This fact is still hidden in the terminology and notation that we have used.

What we write in this note as $g \cdot s$, $gs$, or $g(s)$ is quite often read aloud as "$g$ applied to $s$", mimicking the terminology from functions. Taking a look at the expressions

$$e(s) = s, \qquad (gh)(s) = g(h(s)), \qquad \text{and} \qquad g(g^{-1}(s)) = s = g^{-1}(g(s))$$

shows why this terminology makes sense. In other words, when groups act on sets, *we think of $G$ as a collection of bijective maps $S \to S$, with $e \in G$ the identity map*. This explains why we call the map defining the group "composition" and sometimes write it as $g \circ h$, and it proves why the symmetric group is important—it is merely the set of all bijections on a finite set.

*Be aware*, however, when thinking of group actions this way, that it may happen for two group elements $g, h \in G$ that $g(s) = h(s)$ for all $s \in S$ even though $g \neq h$. In other words, we do *not* require the map $\rho\colon G \to \mathrm{Bij}(S)$ to be *injective*. In practise, this rarely causes problems as long as one keeps this in mind. However, there are cases when there is benefit from limiting yourself to the case when $\rho$ is injective. Such an action is called **faithful**. This term will never show up in this course again, but it becomes increasingly important the further one gets into group theory. We do note one important conceptual statement. We leave the proof to the reader.

**Cayley's Theorem 1.20.** *Any group acts faithfully on at least one set, namely itself. Concretely, there exists an injective group homomorphism $\rho\colon G \hookrightarrow \mathrm{Bij}(G)$ given by $\rho(g)(x) = gx \in G$ (see Example 1.15). Thus any group can be realized as a subgroup of some group of bijections on some set.*

## 1.3 Presentations by generators and relations

This section is a bit on the advanced side and should probably be skipped until the reader feels comfortable about quotient groups. But group presentations (not to be confused with *re*presentations, a quite unrelated concept to be discussed later in this course) is something anyone dealing with groups should be introduced to at some point, and it will keep appearing in future applications of the theory.

Group presentations are a result of the realization that for many groups $G$, we are in possession of a collection $S$ of *generators*, meaning that any group element is a product (possibly empty as in the case of the neutral element) of elements from $S$. To be more formal, if $\langle S \rangle \subset G$ denotes the smallest subgroup of $G$ containing all elements of $S$, then in fact $\langle S \rangle = G$. If the subgroup $\langle S \rangle$ generated by $S$ sounds like a magical concept, simply note that

$$\langle S \rangle = \bigcap_{H \supset S} H$$

satisfies this property, where the intersection runs over all subgroups $H$ containing $S$. This intersection is itself a subgroup, as the reader may verify. Note that this generalizes the subgroup $\langle g \rangle \subset G$ generated by a single element. As an example, Proposition 2.9.14(i) in Lauritzen (2003) shows that $S_n$ is generated by the simple transpositions $S = \{s_1, s_2, \ldots, s_{n-1}\}$, that is, $S_n = \langle S \rangle$.

To describe a group, it is of course not enough to remark a set of elements generating it, because we also need to know how these group elements *interact*, that is, we need to understand the *relations* between them. To formalize this concept, we need a definition.

Let $S$ be any set of elements. The **free group on $S$** is a group $F(S)$ defined in the following rather abstract way: First let $S^{-1}$ be a set disjoint from $S$ and with the same cardinality as $S$, and write the elements as $S^{-1} = \{s^{-1} \mid s \in S\}$; for the time being, $s^{-1}$ is not an actual inverse to $s$ in any traditional way, but is simply an abstract symbol. We shall temporarily call $s$ and $s^{-1}$ "formal inverses". Now define $F(S)$ as the set of all finite sequences

$$s_1 s_2 \cdots s_n$$

of elements $s_i$ belonging to ether $S$ or $S^{-1}$, and where no element $s_i$ is a neighbour to its formal inverse $s_i^{-1}$. We allow the empty sequence, which is just denoted $e$. We refer to such a sequence as a **word** in $S$ (and its formal inverses). The elements $s$ and $s^{-1}$ used to write the words are referred to as **letters**.

To turn $F(S)$ into a group, define the group operation simply by juxtaposition, that is,

$$(s_1 s_2 \cdots s_n)(t_1 t_2 \cdots t_m) = s_1 s_2 \cdots s_n t_1 t_2 \cdots t_m.$$

We use the convention that if any element in this joined sequence is a neighbour to its formal inverse, they annihilate each other. In other words, if $s_n$ and $t_1$ happen to be formal inverses, we simply drop them from the sequence. Then the sequence becomes

$$s_1 s_2 \cdots s_{n-1} t_2 t_3 \cdots t_m.$$

Again, if $s_{n-1}$ and $t_2$ happen to be formal inverses, we again drop them from the sequence. Continuing like this, we eventually arrive at a sequence (possibly empty) that satisfies our defining condition, meaning that no two formal inverses are neighbours.

It is intuitively obvious that this operation is associative, but proving it formally is a tricky combinatorial problem whose answer will be either very technical or very abstract, or both. The proof is beyond the scope of this note, and we refer to mathematical literature; an elegant argument is found in a note distributed during the course *Introduction to Topology*. But once associativity is established, it follows that $F(S)$ is a group, called the *free group* on $S$. We

shall see in a moment that the "free" refers to "free from relations"; indeed, the group consists of a collection of symbols satisfying no relations other than those necessary for it to be a group, e.g. $ss^{-1} = e = s^{-1}s$ for all $s \in S$.

The group $F(S)$ is potentially very "large"; even if $S$ has just two elements, there are very many words that can be written with these. The group $F(S)$ itself is uninteresting and horrible to work with; instead, we are interested in its *quotients groups*. Suppose we would like to create a group $G$ generated by two elements $s, t$, but where we require them to satisfy the slightly ridiculous relation

$$s^{17} = t^{24}. \tag{1.3}$$

There are potentially many different groups satisfying this; indeed, we could simply let $G$ be the trivial group and $s = t = e$. However, using free groups, we can create this group in a way that is as general as possible in the sense that $G$ must only satisfy relations that follow from (1.3). To make this work, define $S$ to be a set with two elements which we conveniently denote $s, t$. Inside the free group $F(S)$, consider the normal subgroup $N(s^{17}t^{-24})$ generated by the element $s^{17}t^{-24}$ (see Remark 1.10); indeed, in the group we want to create, this element should be equal to the identity. Now the group $G = F(S)/N(s^{17}t^{-24})$ is generated by the elements $[s], [t] \in G$ because $F(S)$ is generated by $s, t$, and they satisfy $[s]^{17}[t]^{-24} = [e]$. Having no more use of the free group $F(S)$, we drop the brackets from the notation and arrive at the relation (1.3). Thus $G$ is our desired group.

One can do the same with all kinds of different choices of $S$ and any subset $R \subset F(S)$: The group $G = F(S)/N(R)$ is called the group with **generators** $S$ and **relations** $R$. A choice of $S$ and $R$ such that $G$ is isomorphic to $F(S)/N(R)$ is called a **presentation** of $G$. One often writes $G = \langle S \,|\, R \rangle$ to indicate that $G$ has a presentation with generators $S$ and relations $R$. The group $G$ is called **finitely generated** if it has a presentation $G = \langle S \,|\, R \rangle$ with $S$ finite.

The usefulness of presentations comes from the following fact, which is enlightening and trivial at the same time:

**Proposition 1.21.** *Every group has a presentation.*

***Proof.*** Define the set $S$ simply by $S = G$. There is a surjective group homomorphism $\varphi\colon F(S) \twoheadrightarrow G$, mapping each word in $F(S)$ to the same sequence of elements evaluated in $G$. The map $\varphi$ has a non-trivial (and potentially very large) kernel. By the First Isomorphism Theorem, we have $G \cong F(S)/\mathrm{Ker}(\varphi)$. Therefore, we define $R = \mathrm{Ker}(\varphi)$ and note that $N(R) = R$ since $R$ is already a normal subgroup of $F(S)$. Thus $G \cong F(S)/N(R)$, and $G = \langle S \,|\, R \rangle$ is a presentation of $G$. □

It is only the existence itself that is important; the concrete presentation produced in the proof is essentially useless except in abstract settings, since the group $F(S)$ and the relation set $R$ will usually be extremely large; try, for instance, to see what happens in the case $G = \mathbb{R}$!

**Example 1.22.** The cyclic group $\mathbb{Z}$ is generated by the element 1 with no relations at all, hence $\mathbb{Z} = F(1)$. The cyclic group $\mathbb{Z}/d\mathbb{Z}$ with $d > 0$ is also generated by $[1]$, but with the relation $d[1] = [1] + [1] + \cdots + [1] = [0]$. In other words, we have the presentation $\mathbb{Z}/d\mathbb{Z} = \langle [1] \,|\, d[1] \rangle$. ○

**Example 1.23.** The group $\mathbb{Z}^n$ is generated by the set $S = \{e_1, e_2, \ldots, e_n\}$ of standard basis vectors, subject only to the relation that the group be *Abelian*, which boils down to $e_i + e_j = e_j + e_i$ for all $i, j$. Hence it has a presentation

$$\mathbb{Z}^n = \langle e_1, e_2, \ldots, e_n \,|\, e_i + e_j = e_j + e_i,\ i, j = 1, 2, \ldots, n \rangle.$$

Notice that we have notated the relations using equality signs instead of writing the actual elements of the set $R$ that generates our relations. So formally we should write $(e_i + e_j) - (e_j + e_i)$ instead of $e_i + e_j = e_j + e_i$. The less formal notation is common in mathematical literature and often allows for a more clean description of relations.

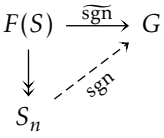
Because the only relations we require are those necessary to make the group Abelian, $\mathbb{Z}^n$ is often called the **free Abelian group** with generators $e_i$. One can define the free Abelian group $\mathbb{Z}^S$ similarly using any (possibly infinite) set $S$ by imposing only the relations that the elements of $S$ *commute*. Free Abelian groups can be used to give presentations of any Abelian group and are extremely important across the mathematical sciences. ○

**Example 1.24.** It can be shown that the symmetric group $S_n$ has a presentation by generators $S = \{s_1, s_2, \ldots, s_{n-1}\}$ and relations

$$s_i^2 = e, \qquad (s_i s_{i+1})^3 = e \qquad s_i s_j = s_j s_i \text{ when } |i - j| > 1.$$

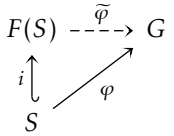
Here these relations hold whenever they make sense, so the middle relation holds for $i = 1, 2, \ldots, n-2$. Thus $S_n$ is isomorphic to $F(S)/N(R)$, where $R \subset S_n$ is the set consisting of $s_i^2$, $(s_i s_{i+1})^3$, and $s_i s_j (s_j s_i)^{-1}$ for all $i$ where this makes sense. We conclude that $S_n$ is finitely generated.

Let us use this presentation to prove the existence of a sign map $\operatorname{sgn}\colon S_n \to \{\pm 1\}$. Since every permutation is a product of simple reflections, the sign homomorphism, if it exists, is uniquely determined by its values on the simple reflections, which is $-1$. We first define a homomorphism $\widetilde{\operatorname{sgn}}\colon F(S) \to \{\pm 1\}$ by declaring that $\widetilde{\operatorname{sgn}}(s_i) = -1$ for all $i$; this means that for every word in $F(S)$ we simply replace every letter by $-1$. This is well-defined since each word in $F(S)$ can be written as a unique combination of letters. The existence of the map sgn defined on $S_n = F(S)/N(R)$ will now follow from the Fundamental Homomorphism Theorem once we show that $\widetilde{\operatorname{sgn}}$ vanishes on $N(R)$, for which it is enough to show that it vanishes on $R$ (why?). But this is clearly the case because each element in $R$ consists of an even number of letters and is thus sent to 1. ○




Actually, the last example contained a hidden universal property of the free group $F(S)$. We state it and leave the proof to the reader:

**Proposition 1.25.** *Suppose that $S$ is a set and $G$ a group, and let $i\colon S \hookrightarrow F(S)$ be the map of sets sending each $s \in S$ to the one-letter word $s \in F(S)$. Then for any map of sets $\varphi\colon S \to G$ there exists a unique group homomorphism $\widetilde{\varphi}\colon F(S) \to G$ satisfying $\varphi = \widetilde{\varphi} \circ i$.*




Be aware, however, that while the technique of generators and relations allows one to create all kinds of exciting groups with exotic properties, it may

(and often will) happen that the relations imposed on the group force it to be *trivial*; this happens if the set $R$ is "too large", so that $N(R)$ becomes all of $F(S)$. There is no general method to determine whether a group given by generators and relations is trivial or not; it usually requires *ad hoc* methods specific to the case at hand.

# Chapter 2

# Rings

## 2.1 Quotient rings

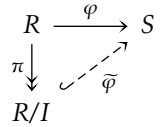
Ideals are the ring-theoretic parallel of normal subgroups, designed with the single purpose of defining quotient rings. When calculating in the quotient ring $R/I$, the ring operations carry over much of the structure from the parent ring $R$ because $[x+y] = [x]+[y]$ and $[x \cdot y] = [x] \cdot [y]$. However, the major difference is that for all $x \in I$ we have $[x] = x + I = I = [0]$, so that the elements of $I$ behave like zero in the quotient ring. Thus we obtain a ring-theoretic parallel to Interpretation 1.1:

**Interpretation 2.1.** *The quotient ring $R/I$ is the ring $R$ where all elements of $I$ are declared to be trivial (i.e. equal to zero). In other words, when calculating in $R/I$, we calculate as in $R$, but ignore elements in $I$.*

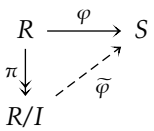
In order to justify that we have chosen the right definition of ideals for this purpose, notice that $(I,+)$ is by definition a subgroup of the Abelian group $(R,+)$ and hence a normal subgroup by Exercise 2.14 in Lauritzen (2003). Thus group theory already provides us with an additive group structure on $R/I$, namely $[x+y] = [x]+[y]$. The additional assumption that $\lambda x \in I$ for all $\lambda \in R$ and $x \in I$ is what allows us to make the multiplication $[x \cdot y] = [x] \cdot [y]$ a well-defined operation as well.

We also obtain ring-theoretic analogues of Norther's Isomorphism Theorems and the Fundamental Homomorphism Theorem. We leave it to the reader to adjust the proofs from the last section.

**Noether's First Isomorphism Theorem 2.2.**
*If $\varphi\colon R \to S$ is a ring homomorphism with kernel $I = \operatorname{Ker}(\varphi)$, then there exists a well-defined map $\widetilde{\varphi}\colon R/I \hookrightarrow S$ given by $\widetilde{\varphi}([x]) = \varphi(x)$ which is an injective ring homomorphism. In terms of the canonical map $\pi\colon R \twoheadrightarrow R/I$, this means that $\varphi = \widetilde{\varphi} \circ \pi$, so that the diagram on the right is commutative.*



**Fundamental Homomorphism Theorem 2.3.**
*If $\varphi\colon R \to S$ is a ring homomorphism and $I \subset R$ is an ideal contained in the kernel* $\mathrm{Ker}(\varphi)$*, then there exists a well-defined map $\widetilde{\varphi}\colon R/I \to S$ given by $\widetilde{\varphi}([x]) = \varphi(x)$ which is a ring homomorphism with kernel* $\mathrm{Ker}(\varphi)/I$*. In terms of the canonical map $\pi\colon R \twoheadrightarrow R/I$, this means that $\varphi = \widetilde{\varphi} \circ \pi$, so that the diagram on the right is commutative.*




Notice that $\mathrm{Ker}(\varphi)$ is in general not a subring (it does not in general contain 1), so $\mathrm{Ker}(\varphi)/I$ is quotient of groups, not of rings. As in the group case, note also that $I \subset \mathrm{Ker}(\varphi)$ is equivalent to requiring that $\varphi$ vanishes on $I$, that is, $\varphi(x) = 0$ for all $x \in I$.

**Advanced Example 2.4.** In this example we develop a ring-theoretic analogue to the Abelianization introduced in Example 1.9. Given a (not necessarily commutative) ring $R$, we look for a way to "make $R$ commutative" while maintaining as much of the structure of $R$ as possible. Namely, we construct a ring $R^{\mathrm{com}}$ that can be thought of abstractly as the ring $R$ where we simply declare that $x \cdot y = y \cdot x$ and calculate as before.

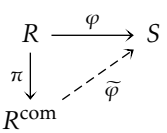
We proceed as before, defining the **commutator** by

$$[x, y] = x \cdot y - y \cdot x, \qquad x, y \in R,$$

and notice that $R$ is commutative if and only if this vanishes for all $x, y$. Then let $I = \langle [x, y] \,|\, x, y \in R \rangle$ be the ideal generated by all possible commutators, and define $R^{\mathrm{com}} = R/I$. Calculating in this quotient, we obtain

$$[0] = [[x, y]] = [x][y] - [y][x]$$

for all $[x], [y] \in R^{\mathrm{com}}$, showing that $R^{\mathrm{com}}$ is commutative. Notice that if $R$ was commutative to begin with, all commutators were already zero, and we have $R = R^{\mathrm{com}}$.




Similarly to the Abelianization of a group, the construction $R^{\mathrm{com}}$ comes with the surjective map $\pi\colon R \twoheadrightarrow R^{\mathrm{com}}$, the canonical map. This has the following universal property: If $\varphi\colon R \to S$ is any ring homomorphism to some commutative ring $S$, there exists a map $\widetilde{\varphi}\colon R^{\mathrm{com}} \to S$ satisfying $\varphi = \widetilde{\varphi} \circ \pi$, that is, the diagram on the right is commutative. Similarly to the group case, this follows from the Fundamental Homomorphism Theorem because the fact that $S$ is commutative means that $\varphi$ vanishes on all commutators. Thus the commutators belong to the kernel $\mathrm{Ker}(\varphi)$, as does the ideal $I$ they generate. We interpret this by saying that $R^{\mathrm{com}}$ is, in a sense, the most general commutative ring with a map $R \to R^{\mathrm{com}}$. ○

**Noether's Second Isomorphism Theorem 2.5.**
*Let $R$ be a ring, $S$ a subring, and $I$ an ideal. Then $S + I$ is a subring of $R$, $S \cap I$ is an ideal in $S$, and there exists an isomorphism of rings*

$$S/(S \cap I) \xrightarrow{\sim} (S + I)/I, \qquad s + (S \cap I) \longmapsto s + I.$$

**Noether's Third Isomorphism Theorem 2.6.**
*Let $R$ be a ring and $I$ an ideal. Then we have:*

(i) *If $S$ is a subring of $R$ containing $I$, then $S/I$ is a subring of $R/I$. Furthermore, every subring of $R/I$ has this form.*

(ii) *If $J$ is an ideal in $R$ containing $I$, then $J/I$ is an ideal in $R/I$. Furthermore, every ideal in $R/I$ has this form. Finally, we have a ring isomorphism*

$$(R/I)/(J/I) \xrightarrow{\sim} R/J, \qquad (r+I)+(J/I) \longmapsto r+J.$$

## 2.2 Fields

Fields are an especially well-behaved class of rings, consisting of the ones in which *division* is always possible. Linear algebra is a very well-behaved subject precisely because we only work over fields; as we shall see in the next section, replacing a field by an arbitrary ring in the definition of a vector space causes everything to become much more complicated and less predictable. At the same time, fields tend to be very complicated objects; it is difficult to construct rings with multiplication and addition maps fitting together in a way that produces a field. Notice how hard finite fields are to construct!

It is important to notice that we de *not* consider the nullring to be a field. This is a matter of convention; the nullring behaves very differently from finite fields. For instance, linear algebra over the nullring breaks down; to name an example, all "vector spaces" over $\{0\}$ are isomorphic to $\{0\}$, hence $\{0\}^n \cong \{0\}$, and there is no notion of dimension. In Lauritzen (2003), the nullring is elegantly excluded in the definition of a field by requiring that $R^* = R \setminus \{0\}$: If $R = \{0\}$, then $R^* = R \neq R \setminus \{0\}$, since any element divides (and is equal to) 1.

Many of the useful properties of fields boil down to the following characterization: Mimicking the notion of a **simple group**, a group without non-trivial normal subgroups, we define a **simple ring** to be a ring without non-trivial ideals, that is, the only ideals are $R$ and $\{0\}$. We have

**Proposition 2.7.** *A simple (commutative) ring is the same thing as a field.*

***Proof.*** If $R$ is a ring and $I \subset R$ a non-zero ideal, then any $x \neq 0$ in $I$ divides 1, hence $1 \in \langle x \rangle \subset I$, and $I = R$. Conversely, if $R$ is simple and $x \neq 0$ is an element, then $\langle x \rangle \neq 0$, hence $\langle x \rangle = R$. In particular, $1 \in \langle x \rangle$, hence $x$ divides 1. □

**Corollary 2.8.** *If $K$ is a field and $R$ an arbitrary ring that is not the nullring, then any ring homomorphism $\varphi\colon K \to R$ is injective. In particular, since the nullring is not a field, any homomorphism between fields is injective.*

***Proof.*** The kernel $\operatorname{Ker}(\varphi)$ is an ideal. It cannot be $R$ since $f(1) = 1 \neq 0$. Hence $\operatorname{Ker}(\varphi) = \{0\}$ and $\varphi$ is injective. □

An injective ring homomorphism is often called an **embedding**, since it provides an isomorphism between the domain and the image; hence the existence of an injective homomorphism $\varphi\colon K \hookrightarrow R$ means that $R$ contains a subring isomorphic to $K$. In other words, we may as well think of $K$ as a subring of $R$

and $\varphi$ as a kind of inclusion. Thus up to identifications, the only maps *from* a field are inclusions (at least when the nullring is excluded).

We note that the characteristic of subring is always the same as the characteristic of the whole ring; this is because a subring is required to contain the element 1. It follows that:

**Corollary 2.9.** *There are no maps between fields of different characteristic.*

We conclude that it is impossible to cook up maps like $\mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/3\mathbb{Z}$, $\mathbb{Q} \to \mathbb{Z}/2\mathbb{Z}$, etc.

## 2.3 Modules over rings

I SHOULD WARN THE READER that this section is advanced and only intended for those feeling comfortable about ring theory and seeking deeper insight into it. Note also that we do *not* assume that rings are commutative in this section.

Similarly to the notion of a group action, rings can also "act" on their surroundings. However, it works quite differently; first of all, the set that a ring acts on must be not just a set, but an *Abelian group*. The further one gets into ring theory, the more obvious it becomes that it is essentially the study of operations on Abelian groups, just like group theory is the study of bijective operations on sets.

An Abelian group $M$ on which a ring $R$ acts is called a *module* over $R$. We have already seen examples of this in linear algebra: A vector space is acted on by a field by means of scalar multiplication. In fact, the definition of an $R$-module is obtained from the definition of a vector space by replacing each occurrence of "field" with "ring".

To be precise, an ***R*-module** is an Abelian group $(M, +)$ together with a map

$$R \times M \longrightarrow M, \qquad (\alpha, m) \longmapsto \alpha \cdot m,$$

called **scalar multiplication**, satisfying the following axioms for all $\alpha, \beta \in R$ and $m, n \in M$:

- $1 \cdot m = m$;
- $\alpha \cdot (\beta \cdot m) = (\alpha\beta) \cdot m$;
- $\alpha \cdot (m + n) = \alpha \cdot m + \alpha \cdot n$;
- $(\alpha + \beta) \cdot m = \alpha \cdot m + \beta \cdot m$.

As with group actions, we often suppress the dot and write $\alpha m$ instead of $\alpha \cdot m$.

Thus a module over a field is the same thing as a vector space over the field (compare to Lauritzen 2003, Definition B.0.9), and it is only for historical and distinctive reasons that we do not call them "vector spaces over $R$". Part of the reason is probably also that modules are not as well-behaved as vector spaces; we may mention that there are many finitely generated $R$-modules that are *not* isomorphic to $R^n$ for some $n$, and that there are some ugly rings (apart from the trivial case of the nullring) where $R^n \cong R^m$ for $n \neq m$! A general introduction

to the theory of modules is given in *Advanced Algebra*, and we only cover the surface here.

A **homomorphism** of $R$-modules (also known as an ***R*-linear map**) is a map $\varphi\colon M \to N$ between modules commuting with scalar multiplication in the natural way, $\varphi(rm) = r\varphi(m)$ for all $r \in R$, $m \in M$. A **submodule** of an $R$-module $M$ is a subgroup $N \subset M$ closed under scalar multiplication, that is, $RN \subset N$. The kernel and image of a module homomorphism are examples of submodules.

**Example 2.10.** Modules over $\mathbb{Z}$ are simply the same as Abelian groups. One way is obvious: A $\mathbb{Z}$-module is by definition an Abelian group. To see that an Abelian group $(M,+)$ also has the structure of $\mathbb{Z}$-module, we simply define the scalar multiplication $\alpha \cdot m$ by letting $\alpha \cdot m$ be $m + \cdots + m$ ($\alpha$ times) if $\alpha \geq 0$ and $\alpha \cdot m = -(m + \cdots + m)$ ($-\alpha$ times) if $\alpha < 0$. ○

**Example 2.11.** The ring $R$ itself becomes an $R$-module with scalar multiplication $R \times R \to R$, $(r,m) \mapsto rm$, the product af $r$ and $m$ in $R$. If $R$ is commutative, then submodules are the same as ideals in $R$. If $R$ is not commutative, one has to distinguish between left ideals (subgroups $I$ satisfying $RI \subset I$), right ideals (subgroups $I$ with $IR \subset I$), and two-sided ideals (subgroups satisfying both), and a submodule is the same as a left ideal. ○

**Advanced Remark 2.12.** What we call a module here should really be called a **left module**. There is a dual notion of **right module** with a scalar multiplication map $M \times R \to M$ satisfying similar axioms. Then the ring $R$ is also a right module, and a submodule is the same thing as a right ideal.

There is essentially nothing new to be gained from passing to right modules, since a right module over $R$ is the same as a left module over $R^{\mathrm{op}}$, the **opposite ring** of $R$, which is the ring obtained from $R$ by reversing the order of multiplication: $x \cdot y$ in $R^{\mathrm{op}}$ is the same as $y \cdot x$ in $R$. △

**Example 2.13.** The finite-dimensional vector space $V = k^n$ over some field $k$ is a module over the ring $\mathrm{M}_n(k)$ of $n \times n$ matrices over $k$ by letting $Av$ be the matrix product of $A \in \mathrm{M}_n(k)$ with $v \in k^n$. ○

**Example 2.14.** For any Abelian group $(M,+)$, there is one canonical ring having $M$ as a module in a very natural way: The ring $R = \mathrm{End}(M)$ of **endomorphisms** of $M$. In general, an endomorphism on an Abelian group is a group homomorphism from the group to itself. The set of all such becomes a ring with the obvious addition map and with multiplication given by composition of maps. This ring is not commutative in general.

To obtain the $R$-module structure on $M$, define a scalar multiplication map by

$$\mathrm{End}(M) \times M \longrightarrow M, \qquad (\varphi, m) \longmapsto \varphi(m),$$

where $\varphi(m)$ simply means the map $\varphi$ applied to $m$. ○

In a sense, the ring $\mathrm{End}(M)$ is the most general ring with a module structure on $M$. To make this precise, recall how in section 1.2 we showed how a group action $G \times S \to S$ is the same as a group homomorphism $G \to \mathrm{Bij}(S)$. Let us try to do the same with modules, the ring-theoretic analogue of actions. Namely,

if $M$ is a module over some ring $R$, define a map $\rho\colon R \to \mathrm{End}(M)$ by mapping each $r \in R$ to the map

$$\rho(r)\colon M \longrightarrow M, \qquad m \longmapsto rm.$$

It then follows from the module axioms that $\rho$ is a ring homomorphism. In other words, we could have equivalently defined an $R$-module to be an Abelian group $M$ together with a ring homomorphism $\rho\colon R \to \mathrm{End}(M)$.

The further one gets into ring theory, particularly its non-commutative part, the clearer it becomes that rings should be understood not as a generalization of number sets, but as a collection of group homomorphisms on some Abelian group. This is due to the fact that Abelian groups are one of the most fundamental objects in mathematics, indispensable across all of the mathematical sciences. In this framework, the objects that rings are meant to generalize turn out not to be the integers or the real numbers, but rings of the form $\mathrm{End}(M)$. Indeed, we can formulate a result analogues to Cayley's Theorem (Theorem 1.20), noting that *any* ring $R$ can be realized as a subring of some endomorphism ring, namely $\mathrm{End}(R)$, the ring of group homomorphisms on $(R,+)$. For some rings, like $R = \mathbb{Z}$, we have $\mathbb{Z} = \mathrm{End}(\mathbb{Z})$, but for most rings the inclusion $R \subset \mathrm{End}(R)$ is strict.

* * *

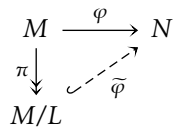
Modules, like groups and rings, allow *quotients*. To make sense of this, note that modules are by definition Abelian groups, hence every subgroup (in particular, every submodule) is normal according to Exercise 2.14 in Lauritzen (2003). Thus we have a group-theoretic quotient $M/L$ for every pair of $R$-modules $M \supset L$. We give this quotient group the structure of an $R$-module the only sober way, defining scalar multiplication by

$$R \times M/L \longrightarrow M/L, \qquad (r, m+L) \longmapsto rm + L.$$

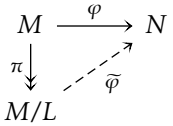
To see that this is well-defined, note that if $m+L = m'+L$, then $m-m' \in L$, hence $r(m-m') \in L$ since $L$ is a submodule. Thus $rm - rm' \in L$, which is equivalent to $rm + L = rm' + L$. Notice that there is no notion of "normal submodule" or "ideal submodule" here; it makes sense to take quotients of *any* module by *any* submodule.

With these definitions, it comes as no surprise that we have module-theoretic versions of the Fundamental Homomorphism Theorem and Noether's Isomorphism Theorems. Again, we leave it to the reader to adjust the proofs from the group-theoretic versions.

**Noether's First Isomorphism Theorem 2.15.**
*If $\varphi\colon M \to N$ is a homomorphism of $R$-modules with kernel $L = \mathrm{Ker}(\varphi)$, then there exists a well-defined map $\widetilde{\varphi}\colon M/L \hookrightarrow N$ given by $\widetilde{\varphi}([m]) = \varphi(m)$ which is an injective module homomorphism. In terms of the canonical map $\pi\colon M \twoheadrightarrow M/L$, this means that $\varphi = \widetilde{\varphi} \circ \pi$, so that the diagram on the right is commutative.*



**Fundamental Homomorphism Theorem 2.16.**
*If $\varphi\colon M \to N$ is a homomorphism of R-modules and L a submodele of M contained in* $\operatorname{Ker}(\varphi)$*, then then exists a well-defined map $\widetilde{\varphi}\colon M/L \to N$ given by $\widetilde{\varphi}([m]) = \varphi(m)$ which is a group homomorphism with kernel* $\operatorname{Ker}(\varphi)/N$*. In terms of the canonical map $\pi\colon M \twoheadrightarrow M/L$, this means that $\varphi = \widetilde{\varphi} \circ \pi$, so that the diagram on the right is commutative.*




**Noether's Second Isomorphism Theorem 2.17.**
*Let M be an R-module with submodules N and L. Then $N + L$ and $N \cap L$ are submodules of M, and there exists an isomorphism of R-modules*

$$N/(N \cap L) \xrightarrow{\sim} (N + L)/L, \qquad n + (N \cap L) \longmapsto n + L.$$

**Noether's Third Isomorphism Theorem 2.18.**
*Let M be an R-module and N a submodule. If L is submodule containing N, then $L/N$ is a submodule of $M/N$. Furthermore, every submodule of $M/N$ has this form. Finally, we have an isomorphism of R-modules*

$$(M/N)/(L/N) \xrightarrow{\sim} M/L, \qquad (m + N) + (L/N) \longmapsto m + L.$$

# Chapter 3

# Polynomials

At the elementary level, we are used to thinking of polynomials as *functions* on some ring, mostly the real and complex numbers. In other words, a real polynomial is a function $f : \mathbb{R} \to \mathbb{R}$ of the form

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

for suitable coefficients $a_i \in \mathbb{R}$. In both the real and complex cases, it can be shown, for instance through differentiation, that the coefficients $a_i$ are uniquely determined. We shall refer to a function defined this way as a **polynomial function** in order to distinguish it from *polynomials*.

For polynomials over arbitrary rings, this definition does not quite provide us with the structure we want. For instance, over the field $\mathbb{F}_2$, there are only two elements in the field and hence a total of four functions $\mathbb{F}_2 \to \mathbb{F}_2$ (it can be checked that all of these do in fact occur as polynomial functions). Thus the coefficients of a polynomial function cannot possibly be unique; indeed, we have $x^n = x$ for all $x \in \mathbb{F}_2$ and $n \geq 1$. However, for our theory of polynomials to make sense, it is necessary to give a definition forcing polynomials to have unique coefficients.

Therefore, we define the ring $R[X]$ of **polynomials** over some (commutative) ring $R$ to be the ring of *formal expression* of the form

$$f = a_0 + a_1 X + a_2 X^2 + a_3 X^3 + \cdots$$

where all coefficients $a_i$ belong to $R$ and all of them are zero but finitely many. By "formal expression", we mean that a polynomial is purely "form", a simple collection of symbols carrying no meaning and where no evaluation is going on. Our $X$ is nothing but an abstract symbol which we manipulate according to rules we specify below. We flatly *declare* that polynomials are equal if and only if all coefficients are equal. We also allow ourselves to omit the infinitely many terms of the form $0X^i$ and simply write

$$f = a_0 + a_1 X + a_2 X^2 + \cdots + a_n X^n.$$

We define addition of polynomials by

$$\begin{aligned}(a_0 + a_1 X + a_2 X^2 + \cdots) + (b_0 + b_1 X + a_2 X^2 + \cdots) \\ = (a_0 + b_0) + (a_1 + b_1)X + (a_2 + b_2)X^2 + \cdots,\end{aligned}$$

summing coefficients of the same degree individually. To define multiplication, we simply define $X^i \cdot X^j = X^{i+j}$ for all $i, j \in \mathbb{N}$ and expand linearly, with the understanding that there is an implicit $X^0$ standing behind the zeroth coefficient. It can then be checked that the $n$th coefficient of the product $f \cdot g$ of $f = a_0 + a_1 X + \cdots$ and $g = b_0 + b_1 X + \cdots$ is

$$\sum_{i+j=n} a_i b_j.$$

With these definitions, $R[X]$ becomes a ring with 0 element $0 = 0 + 0X + 0X^2 + \cdots$ and 1 element $1 + 0X + 0X^2 + \cdots$. In general, we think of $R$ as a subring of $R[X]$ by identifying $r \in R$ with $r + 0X + 0X^2 + \cdots$.

The **degree** of a polynomial $f = a_0 + a_1 X + a_2 X^2 + \cdots$ is the number

$$\deg(f) = \max\{ n \in \mathbb{N} \mid a_n \neq 0 \}.$$

Different authors have different interpretations of what this means for the zero polynomials, where this set is empty. Our course textbook leaves $\deg(0)$ undefined, in accordance with the convention in about half of mathematical literature. The other half defines $\deg(0)$ according to the above formula, with the convention $\max(\varnothing) = -\infty$. By putting $n + (-\infty) = -\infty$ for all $n \in \mathbb{N}$ and $(-\infty) + (-\infty) = -\infty$, a surprising amount of our well-known formulas remain true for *all* polynomials, including $\deg(fg) = \deg(f) + \deg(g)$, and we can often avoid extra considerations for the zero polynomial.

The more concrete definition of polynomials used in Lauritzen (2003) is meant to justify that the notion of "formal expression" used above can be made concrete and is not empty abstract nonsense. There are other, equivalent ways of making things concrete; for instance, the notes Thorup (2007) used for introductory algebra at the University of Copenhagen define polynomials as infinite tuples $f = (a_0, a_1, a_2, \ldots)$ and develop everything from this.

# Bibliography

Lauritzen, N. (2003). *Concrete Abstract Algebra*. Cambridge University Press.
Thorup, A. (2007). *Algebra*. Københavns Univirsitet. URL: http://www.math.ku.dk/noter/filer/alg12.pdf.