3. Modules

In linear algebra, the most important structure is that of a vector space over a field. For commutative algebra it is therefore useful to consider the generalization of this concept to the case where the underlying space of scalars is a commutative ring R instead of a field. The resulting structure is called a *module*; we will introduce and study it in this chapter.

In fact, there is another more subtle reason why modules are very powerful: they unify many other structures that you already know. For example, when you first heard about quotient rings you were probably surprised that in order to obtain a quotient *ring* R/I one needs an *ideal* I of R, i. e. a structure somewhat different from that of a (sub-)ring. In contrast, we will see in Example 3.4 (a) that ideals as well as quotient rings of R are just special cases of modules over R, so that one can deal with both these structures in the same way. Even more unexpectedly, it turns out that modules over certain rings allow a special interpretation: modules over $\mathbb{Z}$ are nothing but Abelian groups, whereas a module over the polynomial ring $K[x]$ over a field K is exactly the same as a K-vector space V together with a linear map $\varphi : V \to V$ (see Examples 3.2 (d) and 3.8, respectively). Consequently, general results on modules will have numerous consequences in many different setups.

So let us now start with the definition of modules. In principle, their theory that we will then quickly discuss in this chapter is entirely analogous to that of vector spaces [G2, Chapters 13 to 18]. However, although many properties just carry over without change, others will turn out to be vastly different. Of course, proofs that are literally the same as for vector spaces will not be repeated here; instead we will just give references to the corresponding well-known linear algebra statements in these cases.

Definition 3.1 (Modules)**.** Let R be a ring. An **R-module** is a set M together with two operations

$$+ : M \times M \to M \qquad \text{and} \qquad \cdot : R \times M \to M$$

(an "addition" in M and a "scalar multiplication" with elements of R) such that for all $m,n \in M$ and $a,b \in R$ we have:

(a) $(M,+)$ is an Abelian group;

(b) $(a+b) \cdot m = a \cdot m + b \cdot m$ and $a \cdot (m+n) = a \cdot m + a \cdot n$;

(c) $(a \cdot b) \cdot m = a \cdot (b \cdot m)$;

(d) $1 \cdot m = m$.

We will also call M a *module over* R, or just a module if the base ring is clear from the context.

Example 3.2.

(a) For a field R, an R-module is by definition exactly the same as an R-vector space [G2, Definition 13.1].

(b) Of course, the zero set $\{0\}$ is a module, which we often simply write as 0.

(c) For $n \in \mathbb{N}_{>0}$ the set $R^n = \{(a_1, \ldots, a_n) : a_1, \ldots, a_n \in R\}$ is an R-module with componentwise addition and scalar multiplication. More generally, for two R-modules M and N the product $M \times N$ with componentwise addition and scalar multiplication is an R-module again.

(d) A $\mathbb{Z}$-module is just the same as an Abelian group. In fact, any $\mathbb{Z}$-module is an Abelian group by definition 3.1 (a), and in any Abelian group $(M,+)$ we can define a multiplication with integers in the usual way by $(-1) \cdot m := -m$ and $a \cdot m := m + \cdots + m$ (a times) for $a \in \mathbb{N}$ and $m \in M$.

(e) Any R-algebra M is also an R-module by Remark 1.24, if we just forget about the possibility to multiply two elements of M.

Definition 3.3 (Submodules, sums, and quotients)**.** Let M be an R-module.

(a) A **submodule** of M is a non-empty subset $N \subset M$ satisfying $m+n \in N$ and $am \in N$ for all $m,n \in N$ and $a \in R$. We write this as $N \leq M$. Of course, N is then an R-module itself, with the same addition and scalar multiplication as in M.

(b) For any subset $S \subset M$ the set

$$\langle S \rangle := \{a_1 m_1 + \cdots + a_n m_n : n \in \mathbb{N}, a_1, \ldots, a_n \in R, m_1, \ldots, m_n \in S\} \quad \subset M$$

of all R-linear combinations of elements of S is the smallest submodule of M that contains S. It is called the submodule **generated by** S. If $S = \{m_1, \ldots, m_n\}$ is finite, we write $\langle S \rangle = \langle \{m_1, \ldots, m_n\} \rangle$ also as $\langle m_1, \ldots, m_n \rangle$. The module M is called **finitely generated** if $M = \langle S \rangle$ for a finite set $S \subset M$.

(c) For submodules $N_1, \ldots, N_n \leq M$ their **sum**

$$N_1 + \cdots + N_n = \{m_1 + \cdots + m_n : m_i \in N_i \text{ for all } i = 1, \ldots, n\}$$

is obviously a submodule of M again. If moreover every element $m \in N_1 + \cdots + N_n$ has a *unique* representation as $m = m_1 + \cdots + m_n$ with $m_i \in N_i$ for all i, we call $N_1 + \cdots + N_n$ a **direct sum** and write it also as $N_1 \oplus \cdots \oplus N_n$.

(d) If $N \leq M$ is a submodule, the set

$$M/N := \{\bar{x} : x \in M\} \quad \text{with} \quad \bar{x} := x + N$$

of equivalence classes modulo N is again a module [G2, Proposition 15.16], the so-called **quotient module** of M modulo N.

Example 3.4.

(a) Let R be a ring. If we consider R itself as an R-module, a submodule of R is by definition the same as an ideal I of R. Moreover, the quotient ring R/I is then by Definition 3.3 (d) an R-module again.

Note that this is the first case where modules and vector spaces behave in a slightly different way: if K is a field then the K-vector space K has no non-trivial subspaces.

(b) The polynomial ring $K[x_1, \ldots, x_n]$ over a field K is finitely generated as a K-algebra (by $x_1, \ldots, x_n$), but not finitely generated as a K-module, i. e. as a K-vector space (the monomials $1, x_1, x_1^2, \ldots$ are linearly independent). So if we use the term "finitely generated" we always have to make sure to specify whether we mean "finitely generated as an algebra" or "finitely generated as a module", as these are two different concepts.

Exercise 3.5. Let N be a submodule of a module M over a ring R. Show:

(a) If N and M/N are finitely generated, then so is M.

(b) If M is finitely generated, then so is M/N.

(c) If M is finitely generated, N need not be finitely generated.

Definition 3.6 (Morphisms)**.** Let M and N be R-modules.

(a) A **morphism** of R-modules (or **R-module homomorphism**, or **R-linear map**) from M to N is a map $\varphi : M \to N$ such that

$$\varphi(m+n) = \varphi(m) + \varphi(n) \qquad \text{and} \qquad \varphi(am) = a\,\varphi(m)$$

for all $m,n \in M$ and $a \in R$. The set of all such morphisms from M to N will be denoted $\mathrm{Hom}_R(M,N)$ or just $\mathrm{Hom}(M,N)$; it is an R-module again with pointwise addition and scalar multiplication.

(b) A morphism $\varphi : M \to N$ of R-modules is called an **isomorphism** if it is bijective. In this case, the inverse map $\varphi^{-1} : N \to M$ is a morphism of R-modules again [G2, Lemma 13.25 (a)]. We call M and N **isomorphic** (written $M \cong N$) if there is an isomorphism between them.

Example 3.7.

(a) For any ideal I in a ring R, the quotient map $\varphi : R \to R/I,\ a \mapsto \overline{a}$ is a surjective R-module homomorphism.

(b) Let M and N be Abelian groups, considered as $\mathbb{Z}$-modules as in Example 3.2 (d). Then a $\mathbb{Z}$-module homomorphism $\varphi : M \to N$ is the same as a homomorphism of Abelian groups, since $\varphi(m+n) = \varphi(m) + \varphi(n)$ already implies $\varphi(am) = a\varphi(m)$ for all $a \in \mathbb{Z}$.

(c) For any R-module M we have $\mathrm{Hom}_R(R,M) \cong M$: the maps

$$M \to \mathrm{Hom}_R(R,M),\ m \mapsto (R \to M, a \mapsto am) \qquad \text{and} \qquad \mathrm{Hom}_R(R,M) \to M,\ \varphi \mapsto \varphi(1)$$

are obviously R-module homomorphisms and inverse to each other. On the other hand, the module $\mathrm{Hom}_R(M,R)$ is in general not isomorphic to M: for the $\mathbb{Z}$-module $\mathbb{Z}_2$ we have $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}_2,\mathbb{Z}) = 0$ by (b), as there are no non-trivial group homomorphisms from $\mathbb{Z}_2$ to $\mathbb{Z}$.

(d) If $N_1,\dots,N_n$ are submodules of an R-module M such that their sum $N_1 \oplus \dots \oplus N_n$ is direct, the morphism

$$N_1 \times \dots \times N_n \to N_1 \oplus \dots \oplus N_n,\ (m_1,\dots,m_n) \mapsto m_1 + \dots + m_n$$

is bijective, and hence an isomorphism. One therefore often uses the notation $N_1 \oplus \dots \oplus N_n$ for $N_1 \times \dots \times N_n$ also in the cases where $N_1,\dots,N_n$ are R-modules that are not necessarily submodules of a given ambient module M.

Example 3.8 (Modules over polynomial rings)**.** Let R be a ring. Then an $R[x]$-module M is the same as an R-module M together with an R-module homomorphism $\varphi : M \to M$:

"$\Rightarrow$" Let M be an $R[x]$-module. Of course, M is then also an R-module. Moreover, multiplication with x has to be R-linear, so $\varphi : M \to M, m \mapsto x \cdot m$ is an R-module homomorphism.

"$\Leftarrow$" If M is an R-module and $\varphi : M \to M$ an R-module homomorphism we can give M the structure of an $R[x]$-module by setting $x \cdot m := \varphi(m)$, or more precisely by defining scalar multiplication by

$$\Big(\sum_{i=0}^{n} a_i x^i\Big) \cdot m := \sum_{i=0}^{n} a_i \varphi^i(m),$$

where φ^i denotes the i-fold composition of φ with itself, and $\varphi^0 := \mathrm{id}_M$.

Remark 3.9 (Images and kernels of morphisms)**.** Let $\varphi : M \to N$ be a homomorphism of R-modules.

(a) For any submodule $M' \leq M$ the image $\varphi(M')$ is a submodule of N [G2, Lemma 13.21 (a)]. In particular, $\varphi(M)$ is a submodule of N, called the **image** of φ.

(b) For any submodule $N' \leq N$ the inverse image $\varphi^{-1}(N')$ is a submodule of M [G2, Lemma 13.21 (b)]. In particular, $\varphi^{-1}(0)$ is a submodule of M, called the **kernel** of φ.

Proposition 3.10 (Isomorphism theorems)**.**

(a) *For any morphism $\varphi : M \to N$ of R-modules there is an isomorphism*

$$M/\ker\varphi \to \mathrm{im}\,\varphi,\ \overline{m} \mapsto \varphi(m).$$

(b) *For R-modules $N' \leq N \leq M$ we have*

$$(M/N')\,/\,(N/N') \ \cong\ M/N.$$

(c) *For two submodules N, N' of an R-module M we have*

$$(N+N')/N' \ \cong\ N/(N \cap N').$$

Proof. The proofs of (a) and (b) are the same as in [G2, Proposition 15.23] and Exercise 1.22, respectively. For (c) note that $N \to (N+N')/N',\ m \mapsto \overline{m}$ is a surjective R-module homomorphism with kernel $N \cap N'$, so the statement follows from (a). □

Exercise 3.11. Let N be a proper submodule of an R-module M. Show that the following statements are equivalent:

(a) There is no submodule P of M with $N \subsetneq P \subsetneq M$.

(b) The module M/N has only the trivial submodules 0 and M/N.

(c) $M/N \cong R/I$ for a maximal ideal $I \trianglelefteq R$.

The concepts so far were all entirely analogous to the case of vector spaces. There are a few constructions however that are only useful for modules due to the existence of non-trivial ideals in the base ring. Let us introduce them now.

Definition 3.12 (IM, module quotients, annihilators)**.** Let M be an R-module.

(a) For an ideal $I \trianglelefteq R$ we set

$$\begin{aligned} IM &:= \langle \{am : a \in I, m \in M\} \rangle \\ &= \{a_1 m_1 + \cdots + a_n m_n : n \in \mathbb{N}, a_1, \ldots, a_n \in I, m_1, \ldots, m_n \in M\}. \end{aligned}$$

Note that IM is a submodule of M, and M/IM is an R/I-module in the obvious way.

(b) For two submodules $N, N' \leq M$ the **module quotient** (not to be confused with the quotient modules of Definition 3.3 (d)) is defined to be

$$N' : N := \{a \in R : aN \subset N'\} \quad \trianglelefteq R.$$

In particular, for $N' = 0$ we obtain the so-called **annihilator**

$$\operatorname{ann} N := \operatorname{ann}_R N := \{a \in R : aN = 0\} \quad \trianglelefteq R$$

of N. The same definition can also be applied to a single element $m \in M$ instead of a submodule N: we then obtain the ideals

$$N' : m := \{a \in R : am \in N'\} \quad \text{and} \quad \operatorname{ann} m := \{a \in R : am = 0\}$$

of R.

Example 3.13.

(a) If M, N, and N' are submodules of the R-module R, i. e. ideals of R by Example 3.4 (a), the product IM and quotient $N' : N$ of Definition 3.12 are exactly the product and quotient of ideals as in Construction 1.1.

(b) If I is an ideal of a ring R then $\operatorname{ann}_R(R/I) = I$.

Let us recall again the linear algebra of vector spaces over a field K. At the point where we are now, i. e. after having studied subspaces and morphisms in general, one usually restricts to finitely generated vector spaces and shows that every such vector space V has a finite basis. This makes V isomorphic to K^n with $n = \dim_K V \in \mathbb{N}$ [G2, Proposition 14.23]. In other words, we can describe vectors by their coordinates with respect to some basis, and linear maps by matrices — which are then easy to deal with.

For a finitely generated module M over a ring R this strategy unfortunately breaks down. Ultimately, the reason for this is that the lack of a division in R means that a linear relation among generators of M cannot necessarily be used to express one of them in terms of the others (so that it can be dropped from the set of generators). For example, the elements $m = 2$ and $n = 3$ in the $\mathbb{Z}$-module $\mathbb{Z}$ satisfy the linear relation $3m - 2n = 0$, but neither is m an integer multiple of n, nor vice versa. So although $\mathbb{Z} = \langle m, n \rangle$ and these two generators are linearly dependent over $\mathbb{Z}$, neither m nor n alone generates $\mathbb{Z}$.

The consequence of this is that a finitely generated module M need not have a linearly independent set of generators. But this means that M is in general not isomorphic to R^n for some $n \in \mathbb{N}$, and thus there is no obvious well-defined notion of dimension. It is in fact easy to find examples for this: $\mathbb{Z}_2$ as a $\mathbb{Z}$-module is certainly not isomorphic to $\mathbb{Z}^n$ for some n.

So essentially we have two choices if we want to continue to carry over our linear algebra results on finitely generated vector spaces to finitely generated modules:

- restrict to R-modules that are of the form R^n for some $n \in \mathbb{N}$; or
- go on with general finitely generated modules, taking care of the fact that generating systems cannot be chosen to be independent, and thus that the coordinates with respect to such systems are no longer unique.

In the rest of this chapter, we will follow both strategies to some extent, and see what they lead to. Let us start by considering finitely generated modules that do admit a basis.

Definition 3.14 (Bases and free modules)**.** Let M be a finitely generated R-module.

(a) We say that a family $(m_1, \dots, m_n)$ of elements of M is a **basis** of M if the R-module homomorphism
$$R^n \to M, \ (a_1, \dots, a_n) \mapsto a_1 m_1 + \dots + a_n m_n$$
is an isomorphism.

(b) If M has a basis, i. e. if it is isomorphic to R^n for some n, it is called a **free** R-module.

Example 3.15. If I is a non-trivial ideal in a ring R then R/I is never a free R-module: there can be no isomorphism
$$\varphi : R^n \to R/I, \ (a_1, \dots, a_n) \mapsto a_1 m_1 + \dots + a_n m_n$$
since in any case $\varphi(0, \dots, 0) = \varphi(a, 0, \dots, 0)$ for every $a \in I$.

Exercise 3.16. Let R be an integral domain. Prove that a non-zero ideal $I \trianglelefteq R$ is a principal ideal if and only if it is a free R-module.

Remark 3.17 (Linear algebra for free modules)**.** Let M and N be finitely generated, free R modules.

(a) Any two bases of M have the same number of elements: assume that we have a basis with n elements, so that $M \cong R^n$ as R-modules. Choose a maximal ideal I of R by Corollary 2.17. Then R/I is a field by Lemma 2.3 (b), and M/IM is an R/I-vector space by Definition 3.12 (a). Its dimension is
$$\dim_{R/I} M/IM = \dim_{R/I} R^n/IR^n = \dim_{R/I} (R/I)^n = n,$$
and so n is uniquely determined by M. We call n the **rank** $\operatorname{rk} M$ of M.

(b) In the same way as for vector spaces, we see that $\operatorname{Hom}_R(R^m, R^n)$ is isomorphic to the R-module $\operatorname{Mat}(n \times m, R)$ of $n \times m$-matrices over R [G2, Proposition 16.11]. Hence, after choosing bases for M and N we also have $\operatorname{Hom}_R(M, N) \cong \operatorname{Mat}(n \times m, R)$ with $m = \operatorname{rk} M$ and $n = \operatorname{rk} N$ [G2, Proposition 16.23].

(c) An R-module homomorphism $\varphi : M \to M$ is an isomorphism if and only if its matrix $A \in \operatorname{Mat}(m \times m, R)$ as in (b) is invertible, i. e. if and only if there is a matrix $A^{-1} \in \operatorname{Mat}(m \times m, R)$ such that $A^{-1}A = AA^{-1} = E$ is the unit matrix. As expected, whether this is the case can be checked with determinants as follows.

(d) For a square matrix $A \in \operatorname{Mat}(m \times m, R)$ the *determinant* $\det A$ is defined in the usual way [G2, Proposition 18.12]. It has all the expected properties; in particular there is an *adjoint matrix* $A^\# \in \operatorname{Mat}(m \times m, R)$ such that $A^\# A = AA^\# = \det A \cdot E$ (namely the matrix with (i, j)-entry $(-1)^{i+j} \det A'_{j,i}$, where $A'_{j,i}$ is obtained from A by deleting row j and column i) [G2, Proposition 18.20 (a)]. With this we can see that A is invertible if and only if $\det A$ is a unit in R:

"$\Rightarrow$" If there is an inverse matrix A^{-1} then $1 = \det E = \det(A^{-1}A) = \det A^{-1} \cdot \det A$, so $\det A$ is a unit in R.

"$\Leftarrow$" If $\det A$ is a unit, we see from the equation $A^\# A = AA^\# = \det A \cdot E$ that $(\det A)^{-1} \cdot A^\#$ is an inverse of A.

05

So all in all finitely generated, free modules behave very much in the same way as vector spaces. However, most modules occurring in practice will not be free — in fact, submodules and quotient modules of free modules, as well as images and kernels of homomorphisms of free modules, will in

general not be free again. So let us now also find out what we can say about more general finitely generated modules.

First of all, the notion of dimension of a vector space, or rank of a free module as in Remark 3.17 (a), is then no longer defined. The following notion of the *length* of a module can often be used to substitute this.

Definition 3.18 (Length of modules)**.** Let M be an R-module.

(a) A **composition series** for M is a finite chain

$$0 = M_0 \subsetneq M_1 \subsetneq M_2 \subsetneq \cdots \subsetneq M_n = M$$

of submodules of M that cannot be refined, i. e. such that there is no submodule N of M with $M_{i-1} \subsetneq N \subsetneq M_i$ for any $i = 1, \ldots, n$. (By Exercise 3.11, this is equivalent to M_i/M_{i-1} having no non-trivial submodules for all i, and to M_i/M_{i-1} being isomorphic to R modulo some maximal ideal for all i).

The number n above will be called the length of the series.

(b) If there is a composition series for M, the shortest length of such a series is called the **length** of M and denoted $l_R(M)$ (in fact, we will see in Exercise 3.19 (b) that then all composition series have this length). Otherwise, we set formally $l_R(M) = \infty$.

If there is no risk of confusion about the base ring, we write $l_R(M)$ also as $l(M)$.

Exercise 3.19. Let M be an R-module of finite length, i. e. an R-module that admits a composition series. Show that:

(a) If $N < M$ is a proper submodule of M then $l(N) < l(M)$.

(b) Every composition series for M has length $l(M)$.

(c) Every chain $0 = M_0 \subsetneq M_1 \subsetneq M_2 \subsetneq \cdots \subsetneq M_n = M$ of submodules of M can be refined to a composition series for M.

Example 3.20.

(a) Let V be a vector space over a field K. If V has finite dimension n, there is a chain

$$0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V$$

of subspaces of V with $\dim_K V_i = i$ for all i. Obviously, this chain cannot be refined. Hence it is a composition series for V, and we conclude by Exercise 3.19 (b) that $l(V) = n = \dim_K V$.

On the other hand, if V has infinite dimension, there are chains of subspaces of V of any length. By Exercise 3.19 this is only possible if $l(V) = \infty$.

So for vector spaces over a field, the length is just the same as the dimension.

(b) There is no statement analogous to (a) for free modules over a ring: already $\mathbb{Z}$ has infinite length over $\mathbb{Z}$, since there are chains

$$0 \subsetneq (2^n) \subsetneq (2^{n-1}) \subsetneq \cdots \subsetneq (2) \subsetneq \mathbb{Z}$$

of ideals in $\mathbb{Z}$ of any length.

(c) Certainly, a module M of finite length must be finitely generated: otherwise there would be infinitely many elements $(m_i)_{i \in \mathbb{N}}$ of M such that all submodules $M_i = \langle m_1, \ldots, m_i \rangle$ are distinct. But then $0 = M_0 \subsetneq M_1 \subsetneq M_2 \subsetneq \cdots$ is an infinite chain of submodules, which by Exercise 3.19 is impossible for modules of finite length.

On the other hand, a finitely generated module need not have finite length, as we have seen in (b). In fact, we will study the relation between the conditions of finite generation and finite length in more detail in Chapter 7.

Exercise 3.21. What are the lengths of $\mathbb{Z}_8$ and $\mathbb{Z}_{12}$ as $\mathbb{Z}$-modules? Can you generalize this statement to compute the length of any quotient ring R/I as an R-module, where I is an ideal in a principal ideal domain R?

Let us now show that the length of modules satisfies the same relations as the dimension of vector spaces when taking sums, intersections, quotients, or images and kernels [G2, Corollary 15.31, Proposition 15.17, and Corollary 15.26].

Proposition 3.22 (Additivity of the length of modules)**.** *For any submodule N of an R-module M we have*

$$l(N)+l(M/N)=l(M).$$

Proof. Let us assume first that $l(M)<\infty$. By Exercise 3.19 (c), the chain $0\leq N\leq M$ can be refined to a composition series

$$0=N_0\subsetneq\cdots\subsetneq N_n=N=M_0\subsetneq\cdots\subsetneq M_m=M \qquad (*)$$

for M, where $l(M)=n+m$ by Exercise 3.19 (b). Of course, the first part of this chain is then a composition series for N, and so $l(N)=n$. Moreover, setting $P_i:=M_i/N$ for $i=1,\dots,m$ we obtain a chain

$$0=P_0\subsetneq\cdots\subsetneq P_m=M/N$$

in which $P_i/P_{i-1}\cong M_i/M_{i-1}$ by Proposition 3.10 (b). As these modules have no non-trivial submodules, we see that the above chain of length m is a composition series for M/N, so that we get the desired result $l(N)+l(M/N)=n+m=l(M)$.

Conversely, if $l(N)$ and $l(M/N)$ are finite, there are composition series

$$0=N_0\subsetneq\cdots\subsetneq N_n=N \quad\text{and}\quad 0=P_0\subsetneq\cdots\subsetneq P_m=M/N$$

for N and M/N, respectively. Setting $M_i:=q^{-1}(P_i)$ with the quotient map $q:M\to M/N$ for all $i=1,\dots,m$, we have $M_i/N=P_i$. So as above, $M_i/M_{i-1}\cong P_i/P_{i-1}$ has no non-trivial submodules, and we obtain a composition series $(*)$ for M. This means that M has finite length as well, and the above argument can be applied to prove the equation $l(N)+l(M/N)=l(M)$ again.

The only remaining case is that both sides of the equation of the proposition are infinite — but then of course the statement is true as well. □

Corollary 3.23.

(a) *For any two submodules M_1,M_2 of an R-module M we have*

$$l(M_1+M_2)+l(M_1\cap M_2)=l(M_1)+l(M_2).$$

(b) *For any R-module homomorphism $\varphi:M\to N$ we have*

$$l(\ker\varphi)+l(\operatorname{im}\varphi)=l(M).$$

Proof.

(a) By Proposition 3.10 (c) we have $(M_1+M_2)/M_2\cong M_1/(M_1\cap M_2)$. Calling this module Q, we obtain by Proposition 3.22

$$l(M_1+M_2)=l(M_2)+l(Q) \qquad\text{and}\qquad l(M_1)=l(M_1\cap M_2)+l(Q).$$

So if $l(M_2)=\infty$ then $l(M_1+M_2)=\infty$, and the statement of the corollary holds. The same is true if $l(M_1\cap M_2)=\infty$ and thus $l(M_1)=\infty$. Otherwise, we obtain

$$l(Q)=l(M_1+M_2)-l(M_2)=l(M_1)-l(M_1\cap M_2),$$

and hence the corollary holds in this case as well.

(b) This is just Proposition 3.22 applied to the homomorphism theorem $M/\ker\varphi\cong\operatorname{im}\varphi$ of Proposition 3.10 (a). □

Remark 3.24. An easy consequence of Corollary 3.23 (b) is that for a homomorphism $\varphi:M\to M$ from a module of finite length to itself we have

$$\varphi\text{ injective}\quad\Leftrightarrow\quad\varphi\text{ surjective}\quad\Leftrightarrow\quad\varphi\text{ bijective}$$

as in [G2, Corollary 15.27], since φ is injective if and only if $l(\ker\varphi)=0$, and surjective if and only if $l(\operatorname{im}\varphi)=l(M)$ (see Exercise 3.19 (a)).

What happens in this statement if we consider a module M that is only finitely generated, but not necessarily of finite length (see Example 3.20 (c))? It is actually easy to see that in this case an injective morphism $\varphi : M \to M$ need not be bijective: the map $\varphi : \mathbb{Z} \to \mathbb{Z}$, $m \mapsto 2m$ is a simple counterexample. In view of this example it is probably surprising that the statement that a surjective map is also bijective still holds — this is what we want to show in Corollary 3.28 below. The main ingredient in the proof is the following generalization of the Cayley-Hamilton theorem from linear algebra.

Proposition 3.25 (Cayley-Hamilton). *Let M be a finitely generated R-module, I an ideal of R, and $\varphi : M \to M$ an R-module homomorphism with $\varphi(M) \subset IM$. Then there is a* monic *polynomial (i. e. its leading coefficient is* 1*)*

$$\chi = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \quad \in R[x]$$

with $a_0, \ldots, a_{n-1} \in I$ and

$$\chi(\varphi) := \varphi^n + a_{n-1}\varphi^{n-1} + \cdots + a_0 \,\mathrm{id} = 0 \quad \in \mathrm{Hom}_R(M, M),$$

where φ^i denotes the i-fold composition of φ with itself.

Proof. Let $m_1, \ldots, m_n$ be generators of M. By assumption we have $\varphi(m_i) \in IM$ for all i, and thus there are $a_{i,j} \in I$ with

$$\varphi(m_i) = \sum_{j=1}^{n} a_{i,j} m_j \quad \text{for all } i = 1, \ldots, n.$$

Considering M as an $R[x]$-module by setting $x \cdot m := \varphi(m)$ for all $m \in M$ as in Example 3.8, we can rewrite this as

$$\sum_{j=1}^{n} (x\delta_{i,j} - a_{i,j}) m_j = 0 \quad \text{with} \quad \delta_{i,j} := \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j \end{cases}$$

for all $i = 1, \ldots, n$. Note that the left hand side of this equation, taken for all i, gives us an element of M^n. If we multiply this from the left with the adjoint matrix of $(x\delta_{i,j} - a_{i,j})_{i,j} \in \mathrm{Mat}(n \times n, R[x])$ as in Remark 3.17 (d), we get

$$\det((x\delta_{i,k} - a_{i,k})_{i,k}) \cdot m_j = 0$$

for all j. So $\chi := \det((x\delta_{i,k} - a_{i,k})_{i,k})$ acts as the zero homomorphism on M, and expanding the determinant shows that the non-leading coefficients of this polynomial lie in fact in I. □

Remark 3.26. If R is a field and thus M a finitely generated vector space, we can only take $I = R$ in Proposition 3.25. In the proof, we can then choose $m_1, \ldots, m_n$ to be a basis of M, so that $(a_{i,j})_{i,j}$ is the matrix of φ with respect to this basis, and χ is the characteristic polynomial of φ [G2, Definitions 19.17 and Remark 19.22]. So in this case the statement of Proposition 3.25 is just the ordinary Cayley-Hamilton theorem for endomorphisms of finite-dimensional vector spaces [G2, Exercise 20.24]. For general rings however, the generators $m_1, \ldots, m_n$ are no longer independent, and so there are several choices for the matrix $(a_{i,j})_{i,j}$.

The following easy consequence of Proposition 3.25 is usually attributed to Nakayama. In fact, there are many versions of Nakayama's lemma in the literature (we will also meet some other closely related statements in Exercise 6.16), but this one is probably one of the strongest. It concerns the construction IM of Definition 3.12 (a) for an ideal I in a ring R and an R-module M. More precisely, let us assume that $M \neq 0$ and $IM = M$. Of course, if R is a field this is only possible if $I = R$, i. e. if $1 \in I$. If R is a general ring however, it does not necessarily follow that $1 \in I$ — but Nakayama's Lemma states that in the case of a finitely generated module M there is at least an element $a \in I$ that acts as the identity on M, i. e. such that $am = m$ for all $m \in M$.

Corollary 3.27 (Nakayama's Lemma). *Let M be a finitely generated R-module, and I an ideal of R with $IM = M$. Then there is an element $a \in I$ with $am = m$ for all $m \in M$.*

Proof. As $M = IM$ we can apply Proposition 3.25 to $\varphi = \mathrm{id}$ and our given ideal I to obtain $a_0, \dots, a_{n-1} \in I$ such that

$$\mathrm{id}^n + a_{n-1}\,\mathrm{id}^{n-1} + \dots + a_0\,\mathrm{id} = (1 + a_{n-1} + \dots + a_0)\,\mathrm{id} = 0 \quad \in \mathrm{Hom}_R(M,M).$$

Setting $a := -a_{n-1} - \dots - a_0 \in I$, this just means that $(1-a)m = 0$, i. e. $am = m$ for all $m \in M$. □

Corollary 3.28. *If M is a finitely generated R-module, any surjective homomorphism $\varphi : M \to M$ is an isomorphism.*

Proof. As in Example 3.8, consider M as an $R[x]$-module by setting $x \cdot m := \varphi(m)$ for all $m \in M$. Then $x \cdot M = \varphi(M) = M$ since φ is surjective, so we can apply Corollary 3.27 with $I = (x)$ to obtain a polynomial $f \in (x)$, i. e. a polynomial $f = a_n x^n + a_{n-1}x^{n-1} + \dots + a_1 x$ without constant coefficient, such that

$$f \cdot m = a_n \varphi^n(m) + a_{n-1}\varphi^{n-1}(m) + \dots + a_1 \varphi(m) = m \quad \text{for all } m \in M.$$

But this means that $\varphi(m) = 0$ implies $m = 0$, and so φ is injective. □

Exercise 3.29. For a prime number $p \in \mathbb{N}$ consider the subring $R = \{\frac{a}{b} : a, b \in \mathbb{Z}, p \not| b\}$ of $\mathbb{Q}$, and let $M = \mathbb{Q}$ as an R-module.

(a) Show that R has exactly one maximal ideal I. Which one?

(In fact, this will be obvious once we have studied localizations — see Example 6.6 and Corollary 6.10.)

(b) For the ideal of (a), prove that $IM = M$, but there is no $a \in I$ with $am = m$ for all $m \in M$.

(c) Find a "small" set of generators for M as an R-module. Can you find a finite one? A minimal one?