Introduction to Modern Algebra II

Math 818 Spring 2023

February 22, 2023

Contents

Chapter 1

Modules

Modules are a generalization of the concept of a vector space to any ring of scalars. But while vector spaces make for a great first example of modules, many of the basic facts we are used to from linear algebra are often a little more subtle over a general ring. These differences are features, not bugs. We will introduce modules, study some general linear algebra, and discuss the differences that make the general theory of modules richer and even more fun.

1.1 Basic assumptions

In this class, all rings have a multiplicative identity, written as 1 or 1_R is we want to emphasize that we are referring to the ring R. This is what some authors call *unital rings*; since for us all rings are unital, we will omit the adjective. Moreover, we will think of 1 as part of the structure of the ring, and thus require it be preserved by all natural constructions. As such, a subring S of R must share the same multiplicative identity with R, meaning $1_R = 1_S$. Moreover, any ring homomorphism must preserve the multiplicative identity. To clear any possible confusion, we include below the relevant definitions.

Definition 1.1. A **ring** is a set R equipped with two binary operations, $+$ and $\cdot$, satisfying:

(1) $(R, +)$ is an abelian group with identity element denoted 0 or 0_R.

(2) The operation $\cdot$ is associative, so that $(R, \cdot)$ is a semigroup.

(3) For all $a, b, c \in R$, we have

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{and} \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

(4) there is a multiplicative identity, written as 1 or 1_R, such that $1 \cdot a = a = a \cdot 1$ for all $a \in R$.

To simplify notation, we will often drop the $\cdot$ when writing the multiplication of two elements, so that ab will mean $a \cdot b$.

Definition 1.2. A ring R is a **commutative ring** if for all $a, b \in R$ we have $a \cdot b = b \cdot a$.

Definition 1.3. A ring R is a **division ring** if $1 \neq 0$ and $R \setminus \{0\}$ is a group under $\cdot$, so every nonzero $r \in R$ has a multiplicative inverse. A **field** is a commutative division ring.

Definition 1.4. A commutative ring R is a **domain**, sometimes called an **integral domain** if it has no zerodivisors: $ab = 0 \Rightarrow a = 0$ or $b = 0$.

For some familiar examples, $\mathrm{M}_n(R)$ (the set of $n \times n$ matrices) is a ring with the usual addition and multiplication of matrices, $\mathbb{Z}$ and $\mathbb{Z}/n$ are commutative rings, $\mathbb{C}$ and $\mathbb{Q}$ are fields, and the real Hamiltonian quaternion ring $\mathbb{H}$ is a division ring.

Definition 1.5. A **ring homomorphism** is a function $f\colon R \to S$ satisfying the following:

- $f(a+b) = f(a) + f(b)$ for all $a, b \in R$.
- $f(ab) = f(a)f(b)$ for all $a, b \in R$.
- $f(1_R) = 1_S$.

Under this definition, the map $f\colon \mathbb{R} \to \mathrm{M}_2(\mathbb{R})$ sending $a \mapsto \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}$ preserves addition and multiplication but not the multiplicative identities, and thus it is not a ring homomorphism.

Exercise 1. For any ring R, there exists a unique homomorphism $\mathbb{Z} \to R$.

Definition 1.6. A subset S of a ring R is a **subring** of R if it is a ring under the same addition and multiplication operations and $1_R = 1_S$.

So under this definition, $2\mathbb{Z}$, the set of even integers, is not a subring of $\mathbb{Z}$; in fact, it is not even a ring, since it does not have a multiplicative identity!

Definition 1.7. Let R be a ring. A subset I of R is an **ideal** if:

- I is nonempty.
- $(I, +)$ is a subgroup of $(R, +)$.
- For every $a \in I$ and every $r \in R$, we have $ra \in I$ and $ar \in I$.

The final property is often called **absorption**. A **left ideal** satisfies only absorption on the left, meaning that we require only that $ra \in I$ for all $r \in R$ and $a \in I$. Similarly, a **right ideal** satisfies only absorption on the right, meaning that $ar \in I$ for all $r \in R$ and $a \in I$.

When R is a commutative ring, the left ideals, right ideals, and ideals over R are all the same. However, if R is not commutative, then these can be very different classes.

One key distinction between unital rings and nonunital rings is that if one requires every ring to have a 1, as we do, then the ideals and subrings of a ring R are very different creatures. In fact, the *only* subring of R that is also an ideal is R itself. The change lies in what constitutes a subring; notice that nothing has changed in the definition of ideal.

Remark 1.8. Every ring R has two **trivial ideals**: R itself and the zero ideal $(0) = \{0\}$.

A **nontrivial ideal** I of R is an ideal that $I \neq R$ and $I \neq (0)$. An ideal I of R is a **proper ideal** if $I \neq R$.

1.2 Modules: definition and examples

Definition 1.9. Let R be a ring with $1 \neq 0$. A **left R-module** is an abelian group $(M, +)$ together with an action $R \times M \to M$ of R on M, written as $(r, m) \mapsto rm$, such that for all $r, s \in R$ and $m, n \in M$ we have the following:

- $(r + s)m = rm + sm$,
- $(rs)m = r(sm)$,
- $r(m + n) = rm + rn$, and
- $1m = m$.

A **right R-module** is an abelian group $(M, +)$ together with an action of R on M, written as $M \times R \to M, (m, r) \mapsto mr$, such that for all $r, s \in R$ and $m, n \in M$ we have

- $m(r + s) = mr + ms$,
- $m(rs) = (mr)s$,
- $(m + n)r = mr + nr$, and
- $m1 = m$.

By default, we will be studying left R-modules. To make the writing less heavy, we will sometimes say **R-module** rather than left R-module whenever there is no ambiguity.

Remark 1.10. If R is a commutative ring, then any left R-module M may be regarded as a right R-module by setting $mr := rm$. Likewise, any right R-module may be regarded as a left R-module. Thus for commutative rings, we just refer to modules, and not left or right modules.

Lemma 1.11 (Arithmetic in modules)**.** *Let R be a ring with $1_R \neq 0_R$ and M be an R-module. Then $0_R m = 0_M$ and $(-1_R)m = -m$ for all $m \in M$.*

Proof. Let $m \in M$. Then

$$0_R m = (0_R + 0_R)m = 0_R m + 0_R m.$$

Since M is an abelian group, the element $0_R m$ has an additive inverse, $-0_R m$, so adding it on both sides we see that

$$0_M = 0_R m.$$

Moreover,

$$m + (-1_R)m = 1_R m + (-1_R)m = (1_R - 1_R)m = 0_R m = 0_M,$$

so $(-1_R)m = -m$. □

Typically, one first encounters modules in an undergraduate linear algebra course: the vector spaces from linear algebra are modules over fields. Later we will see that vector spaces are much simpler modules than modules over other rings. So while one might take linear algebra and vector spaces as an inspiration for what to expect from a module, be warned that this perspective can often be deceiving.

Definition 1.12. Let F be a field. A **vector space** over F is an F-module.

We will see more about vector spaces soon. Note that many of the concepts we will introduce have special names in the case of vector spaces. Here are some other important examples:

Lemma 1.13. *Let M be a set with a binary operation $+$. Then*

(1) M is an abelian group if and only if M is a $\mathbb{Z}$-module.

(2) M is an abelian group such that $nm := \underbrace{m + \cdots + m}_{n \text{ times}} = 0_M$ for all $m \in M$ if and only if M has a $\mathbb{Z}/n$-module structure.

Proof. First, we show 1). If M is a $\mathbb{Z}$-module, then $(M, +)$ is an abelian group by definition of module. Conversely, if $(M, +)$ is an abelian group then there is a unique $\mathbb{Z}$-module structure on M given by the formulas below. The uniqueness of the $\mathbb{Z}$ action follows from the identities below in which the right hand side is determined only by the abelian group structure of M. The various identities follow from the axioms of a module:

$$\begin{cases} i \cdot m = (\underbrace{1 + \cdots + 1}_{i}) \cdot m = \underbrace{1 \cdot m + \cdots + 1 \cdot m}_{i} = \underbrace{m + \cdots + m}_{i} & \text{if } i > 0 \\ 0 \cdot m = 0_M & \\ i \cdot m = -(-i) \cdot m = -(\underbrace{m + \cdots + m}_{-i}) & \text{if } i < 0. \end{cases}$$

We leave it as an exercise to check that this $\mathbb{Z}$-action really satisfies the module axioms.

Now we show 2). If M is a $\mathbb{Z}/n$ module, then $(M, +)$ is an abelian group by definition, and $nm = \underbrace{m + \cdots + m}_{n} = \underbrace{[1]_n \cdot m + \cdots + [1]_n \cdot m}_{n} = [0]_n m = 0_M$.

Conversely, there is a unique $\mathbb{Z}/n$-module structure on M given by the formulas below, which are analogous to the ones above:

$$\begin{cases} [i]_n \cdot m = (\underbrace{[1]_n + \cdots + [1]_n}_{i}) \cdot m = \underbrace{[1]_n \cdot m + \cdots + [1]_n \cdot m}_{i} = \underbrace{m + \cdots + m}_{i} & \text{if } i > 0 \\ 0 \cdot m = 0_M & \\ [i]_n \cdot m = -(-[i]_n) \cdot m = -(\underbrace{m + \cdots + m}_{-i}) & \text{if } i < 0. \end{cases}$$

These formulas are well-defined, meaning they are independent of the choice of representative for $[i]_n$, because of the assumption that $nm = 0_M$. Again checking that this $\mathbb{Z}/n$-action really satisfies the module axioms is left as an exercise. □

The proposition above says in particular that any group of the form

$$G = \mathbb{Z}^\ell \times \mathbb{Z}/d_1 \times \cdots \times \mathbb{Z}/d_m$$

is a $\mathbb{Z}$-module, and if $\ell = 0, m \geqslant 1$ and $d_i \mid n$ for $1 \leqslant i \leqslant m$ then G is also a $\mathbb{Z}/n$-module. In particular, the Klein group is a $\mathbb{Z}/2$-module.

In contrast to vector spaces, for M a module over a ring R, it can happen that $rm = 0$ for some $r \in R$ and $m \in M$ such that $r \neq 0_R$ and $m \neq 0_M$. For example, in the Klein group K_4 viewed as a $\mathbb{Z}$-module we have $2m = 0$ for all $m \in K_4$.

Example 1.14. (1) The trivial R-module is $0 = \{0\}$ with $r0 = 0$ for any $r \in R$.

(2) If R is any ring, then R is a left and right R-module via the action of R on itself given by its internal multiplication.

(3) If I is a left (respectively, right) ideal of a ring R then I is a left (respectively, right) R-module with respect to the action of R on I by internal multiplication.

(4) If R is a subring of a ring S, then S is an R-module with respect to the action of R on S by internal multiplication in S.

(5) If R is a commutative ring with $1 \neq 0$, then $R[x_1, \ldots, x_n]$ is an R-module for any $n \geqslant 1$. This is a special case of (4).

(6) If R is a commutative ring and G is a group, then $R[G]$ is an R-module. This is a special case of (4).

(7) If R is a commutative ring, let $\mathrm{M}_n(R)$ denote set of $n \times n$ matrices with entries in R. Then $\mathrm{M}_n(R)$ is an R-module for $n \geqslant 1$, with the R-action given by multiplying all the entries of the given matrix by the given element of R.

(8) The **free module** over R of rank n is the set

$$R^n = \left\{ \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} \mid r_i \in R, 1 \leqslant i \leqslant n \right\}$$

with componentwise addition and multiplication by elements of R, as follows:

$$\begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} + \begin{bmatrix} r_1' \\ \vdots \\ r_n' \end{bmatrix} = \begin{bmatrix} r_1 + r_1' \\ \vdots \\ r_n + r_n' \end{bmatrix} \text{ and } r \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} = \begin{bmatrix} rr_1 \\ \vdots \\ rr_n \end{bmatrix}.$$

We will often write the elements of R^n as n-tuples $(r_1, \ldots, r_n)$ instead. Notice that R is the free R-module of rank 1.

(9) More generally, given a collection of R-modules $\{A_i\}$, the abelian group

$$\bigoplus_i A_i = \{(a_i)_i \mid a_i \in A_i, a_i = 0 \text{ for all } i \text{ but finitely many}\}$$

is an R-module with the R-action $r(a_i) := (ra_i)$.

1.3 Submodules and restriction of scalars

Definition 1.15. Let R be a ring and let M be a left R-module. An **R-submodule** of M is a subgroup N of M satisfying $rn \in N$ for all $r \in R$ and $n \in N$.

The submodules of an R-module M are precisely the subsets of M which are modules in their own right, via the same R-action as we are considering for M.

Exercise 2. Show that if N is a submodule of M, then N is an R-module via the restriction of the action of R on M to the subset N.

Example 1.16. Every R-module M has two **trivial submodules**: M itself and the **zero module** $0 = \{0_M\}$. A submodule N of M is **nontrivial** if $N \neq M$ and $N \neq 0$.

Lemma 1.17 (One-step test for submodules)**.** *Let R be a ring with $1 \neq 0$ and let M be a left R-module. A nonempty subset N of M is an R-submodule of M if and only if $rn + n' \in N$ for all $r \in R$ and $n, n' \in N$.*

Proof. The One-step Test for subgroups says that if for all $n, n' \in N$ we have $n' - n \in N$, then N is a subgroup of M. By Lemma 1.11, by taking $r = -1$ we get $rn + n' = n' - n$, and by assumption this is an element of N. Therefore, N is a subgroup of M. As a consequence, $0_M \in N$. By taking $n' = 0_M$, we see that for all $n \in N$ and all $r \in R$ we have $rn = rn + n' \in N$, and thus we can now conclude that N is a submodule of M. □

Example 1.18. Let R be a ring and let M be a subset of R. Then M is a left (respectively, right) R-submodule of R if and only if M is a left (respectively, right) ideal of R.

Exercise 3. Let R be a ring and let A and B be submodules of an R-module M. Then the **sum** of A and B,

$$A + B := \{a + b \mid a \in A, b \in B\},$$

and $A \cap B$ are both R-submodules of M.

Exercise 4. Let R be a commutative ring with $1 \neq 0$, let I be an ideal of R and let M be an R-module. Show that

$$IM := \left\{ \sum_{k=1}^{n} j_k m_k \mid n \geqslant 0, j_k \in I, m_k \in M \text{ for } 1 \leqslant k \leqslant n \right\}$$

is a submodule of M.

Example 1.19. When R is a field, the submodules of a vector space V are precisely the subspaces of V. When $R = \mathbb{Z}$, then the class of R-modules is simply the class of all abelian groups, by Lemma 1.13. The submodules of a $\mathbb{Z}$-module M coincide with the subgroups of the abelian group M.

Definition 1.20. Let R be a ring with $1 \neq 0$ and let M be an R-module. Given elements $m_1, \ldots, m_n \in M$, the **submodule generated by** $m_1, \ldots, m_n$ is the subset of M given by

$$Rm_1 + \cdots + Rm_n := \{r_1 m_1 + \cdots + r_n m_n \mid r_1, \ldots, r_n \in R\}.$$

Exercise 5. Let R be a ring with $1 \neq 0$ and M be an R-module. Given $m_1, \ldots, m_n \in M$, the submodule generated by $m_1, \ldots, m_n$ is indeed a submodule of M. Moreover, this is the smallest submodule of M that contains $m_1, \ldots, m_n$, meaning that every submodule of M containing $m_1, \ldots, m_n$ must also contain $Rm_1 + \cdots + Rm_n$.

Definition 1.21. Let R be a ring with $1 \neq 0$. An R-module M is **cyclic** if there exists an element $m \in M$ such that

$$M = Rm := \{rm \mid r \in R\}.$$

Given an R-module M, the ring R is often referred to as the **ring of scalars**, by analogy to the vector space case. Given an action of a ring of scalars on a module, we can sometimes produce an action of a different ring of scalars on the same set, producing a new module structure.

Lemma 1.22 (Restriction of scalars)**.** *Let $\phi\colon R \to S$ be a ring homomorphism. Any left S-module M may be regarded via* **restriction of scalars** *as a left R-module with R-action defined by $rm := \phi(r)m$ for any $m \in M$. In particular, if R is a subring of a ring S, then any left S-module M may be regarded via restriction of scalars as a left R-module with R-action defined by the action of the elements of R viewed as elements of S.*

Proof. Let $r, s \in R$ and $m, n \in M$. One checks that the axioms in the definition of a module hold for the given action using properties of ring homomorphisms. For example:

$$(r+s)m = \phi(r+s)m = (\phi(r) + \phi(s))m = \phi(r)m + \phi(s)m = rm + sm.$$

The remaining properties are left as an exercise. □

Note that the second module structure on M obtained via restriction of scalars is induced by the original module structure, so the two are related. In general, one can give different module structures on the same abelian group over different, possibly unrelated, rings.

Example 1.23. If I is an ideal of a ring R, applying restriction of scalars along the quotient homomorphism $q\colon R \to R/I$ tells us that any left R/I-module is also a left R-module. In particular, applying this to the R/I-module R/I makes R/I a left and right R-module by restriction of scalars along the quotient homomorphism. Thus the R-action on R/I is given by

$$r \cdot (a + I) := ra + I.$$

Example 1.24. Given any ring R there exists a unique ring homomorphism $\mathbb{Z} \to R$, by Exercise 1. Thus any R-module can be given the structure of a $\mathbb{Z}$-module by restriction of scalars along this unique map. Note also that a module over any ring is in particular an abelian group, so we can always regard any R-module as a $\mathbb{Z}$-module by forgetting the R-action and focusing only on the abelian group structure. These two constructions – the restriction of scalars to $\mathbb{Z}$ and the *forgetful functor*[1] – actually coincide.

The next example explains why restriction of scalars is called a *restriction*.

Example 1.25. Let R be a subring of S, and let $i\colon R \to S$ be the inclusion map, which must by definition be a ring homomorphism. Applying restriction of scalars to an S-module M via i is the same as simply *restricting* our scalars to the elements of R.

[1] This is a concrete abstract nonsense construction that we will discuss in Homological Algebra next Fall.

1.4 Module homomorphisms and isomorphisms

Definition 1.26. Given R-modules M and N, an **R-module homomorphism** from M to N is a function $f\colon M \to N$ such that for all $r \in R$ and $m, n \in M$ we have

- $f(m+n) = f(m) + f(n)$
- $f(rm) = rf(m)$.

Remark 1.27. The condition $f(m+n) = f(m) + f(n)$ says that f is a homomorphism of abelian groups, and the condition $f(rm) = rf(m)$ says that f is R-linear, meaning that it preserves the R-action. Since f is a homomorphism of abelian groups, it follows that $f(0) = 0$ must hold.

Definition 1.28. Let M and N be vector spaces over a field F. A **linear transformation** from M to N is an F-module homomorphism $M \to N$.

Example 1.29. Let R be a commutative ring and M be an R-module. For each $r \in R$, the multiplication map $\mu_r : M \to M$ given by $\mu_r(m) = rm$ is a homomorphism of R-modules: indeed, by the definition of R-module we have

$$\mu_r(m+n) = r(m+n) = rm + rn = \mu_r(m) + \mu_r(n),$$

and

$$\mu_r(sm) = r(sm) = (rs)m = (sr)m = s(rm) = s\mu_r(m).$$

Definition 1.30. An R-module homomorphism $h\colon M \to N$ is an **R-module isomorphism** if there is an R-module homomorphism $g : N \to M$ such that $h \circ g = \mathrm{id}_N$ and $g \circ h = \mathrm{id}_M$. We say M and N are **isomorphic**, denoted $M \cong N$, if there exists an isomorphism $M \to N$.

To check that an R-module homomorphism $f\colon M \to N$ is an isomorphism, it is sufficient to check that it is bijective.

Exercise 6. Let $f\colon M \to N$ be a homomorphism of R-modules. Show that if f is bijective, then its set-theoretic inverse $f^{-1}\colon N \to M$ is an R-module homomorphism. Therefore, every bijective homomorphism of R-modules is an isomorphism.

One should think of a module isomorphism as a relabelling of the names of the elements of the module. If two modules are isomorphic, that means that they are *essentially the same*, up to renaming the elements.

Definition 1.31. Let $f\colon M \to N$ be a homomorphism of R-modules. The **kernel** of f is

$$\ker(f) := \{m \in M \mid f(m) = 0\}.$$

The **image** of f, denoted $\mathrm{im}(f)$ or $f(M)$, is

$$\mathrm{im}(f) := \{f(m) \mid m \in M\}.$$

Exercise 7. Let R be a ring with $1 \neq 0$, let M be an R-module, and let N be an R-submodule of M. Then the inclusion map $i\colon N \to M$ is an R-module homomorphism.

Exercise 8. If $f\colon M \to N$ is an R-module homomorphism, then $\ker(h)$ is an R-submodule of M and $\mathrm{im}(f)$ is an R-submodule of N.

Definition 1.32. Let R be a ring and let M and N be R-modules. Then $\mathrm{Hom}_R(M,N)$ denotes the set of all R-module homomorphisms from M to N, and $\mathrm{End}_R(M)$ denotes the set $\mathrm{Hom}_R(M,M)$. We call $\mathrm{End}(M)$ the **endomorphism ring** of M, and elements of $\mathrm{End}(M)$ are called **endomorphisms** of M.

The endomorphism ring of an R-module M is called that because it *is* a ring, with multiplication given by composition of endomorphisms, 0 given by the zero map (the constant equal to 0), and 1 given by the identity map. However, two homomorphisms from M to N are not composable unless $M = N$, so $\mathrm{Hom}_R(M,N)$ is not a ring.

When R is commutative, $\mathrm{Hom}_R(M,N)$ is, however, an R-module; let us describe its R-module structure. Given $f, g \in \mathrm{Hom}_R(M,N)$, $f+g$ is the map defined by

$$(f+g)(m) := f(m) + g(m),$$

and given $r \in R$ and $f \in \mathrm{Hom}_R(M,N)$, $r \cdot f$ is the R-module homomorphism defined by

$$(r \cdot f)(m) := r \cdot f(m) = f(rm).$$

The zero element of $\mathrm{Hom}_R(M,N)$ is the **zero map**, the constant equal to 0_N.

Lemma 1.33. *Let M and N be R-modules over a commutative ring R. Then the addition and multiplication by scalars defined above make $\mathrm{Hom}_R(M,N)$ an R-module.*

Proof. There are many things to check, including:

- The addition and the R-action are both well-defined: given $f, g \in \mathrm{Hom}_R(M,N)$ and $r \in R$, we always have $f+g, rf \in \mathrm{Hom}_R(M,N)$.
- The axioms of an R-module are satisfied for $\mathrm{Hom}_R(M,N)$.

We leave the details as exercises. □

We will see later that for an n-dimensional vector space V over a field F, there is an isomorphism of vector spaces $\mathrm{End}_F(V) \cong M_n(F)$. This says that every linear transformation $T : V \to V$ corresponds to some $n \times n$ matrix. However, the story for general R-modules is a lot more complicated.

Lemma 1.34. *For any commutative ring R with $1 \neq 0$ and any R-module M there is an isomorphism of R-modules $\mathrm{Hom}_R(R,M) \cong M$.*

Before we write a formal proof, it helps to think about *why* this theorem is true. What does it mean to give an R-module homomorphism $f : R \to M$? More precisely, what information do we need to determine such an f? Do we need to be given the values of $f(r)$ for every $r \in R$? Since f is a homomorphism of R-modules, for any $r \in R$ we have

$$f(r) = f(r \cdot 1) = rf(1),$$

so the value of $f(1)$ *completely determines* which R-module homomorphism we are talking about. On the other hand, we can choose *any* $m \in M$ to be the image of 1, since thanks to the axioms for modules, the function

$$f(r) := rm$$

is a well-defined R-module homomorphism for any $m \in M$. In summary, to give an R-module homomorphism $R \to M$ is the same as choosing an element $m \in M$, and $\mathrm{Hom}_R(R, M) \cong M$.

Proof. Let $f\colon M \to \mathrm{Hom}_R(R, M)$ be given for each $m \in M$ by $f(m) = \phi_m$ where ϕ_m is the map defined by $\phi_m(r) = rm$ for all $r \in R$. Now we have many things to check:

- f is well-defined, meaning that for any $m \in M$, its image $f(m) = \phi_m$ is an element of $\mathrm{Hom}_R(R, M)$, since

$$\phi_m(r_1 + r_2) = (r_1 + r_2)m = r_1m + r_2m = \phi_m(r_1) + \phi_m(r_2)$$

$$\phi_m(r_1r_2) = (r_1r_2)m = r_1(r_2m) = r_1\phi_m(r_2)$$

  for all $r_1, r_2 \in R$.

- f is an R-module homomorphism, since

$$\phi_{m_1+m_2}(r) = r(m_1 + m_2) = rm_1 + rm_2 = \phi_{m_1}(r) + \phi_{m_2}(r)$$

$$\phi_{r'm}(r) = r(r'm) = (rr')m = r'(rm) = r'\phi_m(r)$$

- f is injective, since $\phi_m = \phi_{m'}$ implies in particular that $\phi_m(1_R) = \phi_{m'}(1_R)$, which by definition of ϕ_- means that $m = m'$.

- f is surjective, since for $\psi \in \mathrm{Hom}_R(R, M)$ we have $\psi(r) = \psi(r1_R) = r\psi(1_R)$ for all $r \in R$, so $\psi = \phi_{\psi(1_R)}$.

This shows that f is an R-module isomorphism. □

Definition 1.35. Let R be a commutative ring with $1_R \neq 0_R$. An **R-algebra** is a ring A with $1_A \neq 0_A$ together with a ring homomorphism $f\colon R \to A$ such that $f(R)$ is contained in the center of A.

Given an R-algebra A, the R-algebra structure on A induces a natural R-module structure: given elements $r \in R$ and $a \in A$, the R-action is defined by

$$r \cdot a := f(r)a,$$

where the product on the right is the multiplication in A. Similarly, we get a natural right R-module structure on A, and since by definition $f(R)$ is contained in the center of A, we obtain what is called a *balanced bimodule* structure on A. We will discuss these further in Homological Algebra next Fall.

Example 1.36. Let R be a commutative ring with $1_R \neq 0_R$. The ring $R[x_1, \ldots, x_n]$ together with the inclusion map $R \hookrightarrow R[x_1, \ldots, x_n]$ is an R-algebra. More generally, any quotient of $R[x_1, \ldots, x_n]$ is an R-algebra.

The ring of matrices $M_n(R)$ with the homomorphism $r \mapsto rI_n$ is also an R-algebra, as is the group ring $R[G]$ for any group G with the inclusion of R into $R[G]$ given by $r \mapsto re_G$.

Lemma 1.37. *Let R be a commutative ring with $1 \neq 0$ and let M be an R-module. Then $\mathrm{End}_R(M)$ is an R-algebra, with addition and R-action defined as above, and multiplication defined by composition $(fg)(m) = f(g(m))$ for all $f, g \in \mathrm{End}_R(M)$ and all $m \in M$.*

Proof. There are many things to check here, including that:

- The axioms of a (unital) ring are satisfied for $\mathrm{End}_R(M)$.
- There is a ring homomorphism $f\colon R \to \mathrm{End}_R(M)$ such that $f(1_R) = 1_{\mathrm{End}_R(M)} = \mathrm{id}_M$ and $f(R) \subseteq Z(\mathrm{End}_R(M))$.

We will just check the last item and leave the others as exercises. Define $f : R \to \mathrm{End}_R(M)$ by $f(r) = r\,\mathrm{id}_M$. Notice that this is the map μ_r from Example 1.29. Then

$$f(r+s) = (r+s)\,\mathrm{id}_M = r\,\mathrm{id}_M + s\,\mathrm{id}_M = f(r) + f(s)$$

and

$$f(rs) = (rs)\,\mathrm{id}_M = (r\,\mathrm{id}_M) \circ (s\,\mathrm{id}_M) = f(r)f(s)$$

show that f is a ring homomorphism. Moreover, $\mathrm{id}_M \in Z(\mathrm{End}_R(M))$, and once can check easily that $\mu_r \in \mathrm{End}_R(M)$: given any other $g \in \mathrm{End}_R(M)$, and any $m \in M$, since g is R-linear we have

$$(g \circ \mu_r)(m) = g(\mu_r(m)) = g(rm) = rg(m) = (\mu_r \circ g)(m).$$

This shows that $f(R) \subseteq \mathrm{End}_R(M)$. □

Remark 1.38. Let R be a commutative ring with $1 \neq 0$ and let M be an R-module. Then M is also an $\mathrm{End}_R(M)$-module with the action $\phi m = \phi(m)$ for any $\phi \in \mathrm{End}_R(M)$, $m \in M$.

Definition 1.39. Let R be a ring, let M be an R-module, and let N be a submodule of M. The quotient module M/N is the quotient group M/N with R action defined by

$$r(m+N) := rm + N$$

for all $r \in R$ and $m + N \in M/N$.

Lemma 1.40. *Let R be a ring, let M be an R-module, and let N be a submodule of M. The quotient module M/N is an R-module, and the quotient map $q\colon M \to M/N$ is an R-module homomorphism with kernel $\ker(q) = N$.*

Proof. Among the many things to check here, we will only check the well-definedness of the R-action on M, and leave the others as exercises. To check well-definedness, consider $m + N = m' + N$. Then $m - m' \in N$, so $r(m - m') \in N$ by the definition of submodule. This gives that $rm - rm' \in N$, hence $rm + N = rm' + N$. □

Definition 1.41. Given an R-module M and a submodule N of M, the map $q\colon M \to M/N$ is the **canonical quotient map**, or simply the canonical map from M to N.

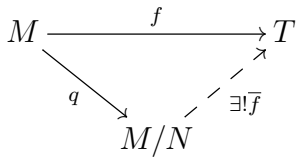
Example 1.42. If R is a field, quotient modules are the same thing as quotient vector spaces. When $R = \mathbb{Z}$, recall that $\mathbb{Z}$-modules are the same as abelian groups, by Lemma 1.13. Quotients of $\mathbb{Z}$-modules coincide with quotients of abelian groups.

Theorem 1.43. *Let N be a submodule of M, let T be an R-module, and let $f : M \to T$ be an R-module homomorphism. If $N \subseteq \ker f$, then the function*

$$\begin{aligned} M/N &\xrightarrow{\overline{f}} T \\ m + N &\longmapsto f(m) \end{aligned}$$

is a well-defined R-module homomorphism. In fact, $\overline{f} : M/N \to T$ is the unique R-module homomorphism such that $\overline{f} \circ q = f$, where $q\colon M \to M/N$ denotes the canonical map.

We can represent this in a more visual way by saying that $\overline{f}$ is the unique R-module homomorphism that makes the diagram



commute.

Proof. By 817, we already know that $\overline{f}$ is a well-defined homomorphism of groups under $+$ and that it is the unique one such that $\overline{f} \circ q = f$. It remains only to show $\overline{f}$ is an R-linear map:

$$\overline{f}(r(m+N)) = \overline{f}(rm+N) = f(rm) = rf(m) = r\overline{f}(m+N).$$

where the third equation uses that f preserves scaling. □

Theorem 1.44 (First Isomorphism Theorem)**.** *Let N be an R-module and let $h : M \to N$ be an R-module homomorphism. Then $\ker(h)$ is a submodule of M and there is an R-module isomorphism $M/\ker(h) \cong \operatorname{im}(h)$.*

Proof. If we forget the multiplication by scalars in R, by the First Isomorphism Theorem for Groups, we know that there is an isomorphism of abelian groups under $+$, given by

$$\begin{aligned} \overline{h} : M/\ker(h) &\xrightarrow{\cong} \operatorname{im}(h) \\ m + \ker(f) &\longmapsto h(m). \end{aligned}$$

It remains only to show this map preserves multiplication by scalars. And indeed:

$$\begin{aligned} \overline{h}(r(m+\ker(h))) &= \overline{h}(rm + \ker(h)) && \text{by definition of the } R\text{-action on } M/\ker(h) \\ &= h(rm) && \text{by definition of } \overline{h} \\ &= rh(m) && \text{since } h \text{ is an } R\text{-module homomorphism} \\ &= r\overline{h}(m + \ker h) && \text{by definition of } h. \end{aligned}$$

Theorem 1.45 (Second Isomorphism Theorem)**.** *Let A and B be submodules of M, and let $A+B=\{a+b \mid a\in A, b\in B\}$. Then $A+B$ is a submodule of M, $A\cap B$ is a submodule of A, and there is an R-module isomorphism $(A+B)/B\cong A/(A\cap B)$.*

Proof. By Exercise 3, $A+B$ and $A\cap B$ are submodules of M. By the Second Isomorphism Theorem for Groups, there is an isomorphism of abelian groups

$$\begin{aligned} h\colon A/(A\cap B) &\xrightarrow{\cong} (A+B)/B \\ a+(A\cap B) &\longmapsto a+B \end{aligned}$$

It remains only to show h preserves multiplication by scalars:

$$h(r(a+(A\cap B))) = h(ra+A\cap B) = ra+B = r(a+B) = rh(a+(A\cap B)). \qquad \square$$

Theorem 1.46 (Third Isomorphism Theorem)**.** *Let A and B be submodules of M with $A\subseteq B$. Then there is an R-module isomorphism $(M/A)/(B/A)\cong M/B$.*

Proof. From 817, we know that B/A is a subgroup of M/A under $+$. Given $r\in R$ and $b+A\in B/A$ we have $r(b+A)=rb+A$ which belongs to B/A since $rb\in B$. This proves B/A is a submodule of M/A. By the Third Isomorphism Theorem for Groups, there is an isomorphism of abelian groups

$$\begin{aligned} (M/A)/(B/A) &\longrightarrow M/B \\ (m+A)+B/A &\longmapsto m+B \end{aligned}$$

and it remains only to show this map is R-linear:

$$\begin{aligned} h(r((m+A)+B/A)) =& h(r(m+A)+B/A) = h((rm+A)+B/A) \\ &= rm+B = r(m+B) \\ &= rh((m+A)+B/A). \qquad \square \end{aligned}$$

Theorem 1.47 (Lattice Isomorphism Theorem)**.** *Let R be a ring, let N be a R-submodule of an R-module M, and let $q\colon M\to M/N$ be the quotient map. Then the function*

$$\begin{aligned} \{R\text{-submodules of } M \text{ containing } N\} &\xrightarrow{\Psi} \{R\text{-submodules of } M/N\} \\ K &\longmapsto K/N \end{aligned}$$

is a bijection, with inverse defined by

$$\Psi^{-1}(T) := q^{-1}(T) = \{a\in M \mid a+N\in T\}$$

for each R-submodule T of M/N. Moreover, Ψ and Ψ^{-1} preserve sums and intersections of submodules.

Proof. From 817, we know there is a bijection between the set of subgroups of M and that contain N and subgroups of the quotient group M/N, given by the same map Ψ. We just need to prove that these maps send submodules to submodules. If K is a submodule of M containing N, then by the Third Isomorphism Theorem we know that K/N is a submodule of M/N. If T is a submodule of M/N, then $\pi^{-1}(T)$ is an abelian group, by 817. For $r \in R$ and $m \in \pi^{-1}(T)$, we have $\pi(m) \in T$, and hence $\pi(rm) = r\pi(m) \in T$ too, since T is a submodule. This proves $\pi^{-1}(T)$ is a submodule. □

We come to a very important class of examples which will help us study linear transformations using module theory.

Lemma 1.48 ($F[x]$-modules)**.** *Let F be a field. There is a bijection*

$$\{V \text{ an } F[x]\text{-module}\} \longleftrightarrow \{V \text{ an } F\text{-vector space and } T \in \mathrm{End}_F(V)\}.$$

Proof. If V is an $F[x]$ module then V is an F-vector space by restriction of scalars along the inclusion $F \hookrightarrow F[x]$. Let $T : V \to V$ be defined by $T(v) = xv$. To show that $T \in \mathrm{End}_F(V)$, note that for any $c \in F$ and $v, v_1, v_2 \in V$ the axioms of the $F[x]$-module give us

$$T(v_1 + v_2) = x(v_1 + v_2) = xv_1 + xv_2 = T(v_1) + T(v_2) \text{ and } T(cv) = x(cv) = c(xv).$$

Conversely, let V be an F-vector space and $T \in \mathrm{End}_F(V)$. We claim that the action of of $F[x]$ on V given by

$$f(x)v = (f(T))(v)$$

satisfies the axioms for a module (exercise!). Alternatively, we can explain this module structure in a more conceptual way, as follows. Consider the evaluation homomorphism $\varphi : F[x] \to \mathrm{End}_F(V), \quad \varphi(f(x)) = f(T)$. Since V is an $\mathrm{End}_F(V)$-module by Remark 1.38, then V is also an $F[x]$-module by restriction of scalars along ϕ; the $F[x]$ action is the one we described above:

$$f(x)v = \varphi(f)(v) = (f(T))(v)$$

Finally, one can check that the two constructions above are inverse to each other. □

Notation 1.49. We shall denote the $F[x]$-module structure on an F-vector space V induced by $T \in \mathrm{End}_F(V)$ by V_T.

Example 1.50. The proposition above says that if we fix an F-vector space V then any linear transformation T gives a different $F[x]$ module structure on V. For example,

- for $T = 0$ the $F[x]$ module V_0 carries an action given by scaling by the constant coefficient of f, that is if $f(x) = a^n x^n + \cdots + a_0$ then
$$f(x)v = (f(0))v = a_0 v \text{ for all } f \in F[x].$$
- for T the "shift operator" that takes $T(e_i) = e_{i-1}$, where e_i is the i-th standard basis vector, the $F[x]$ module V_T is has the action $x^m \begin{bmatrix} v_1 \\ \vdots \\ v_{n-m} \\ v_{n-m+1} \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} v_{m+1} \\ \vdots \\ v_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

1.5 Module generators, bases and free modules

Definition 1.51. Let M be an R-module. A **linear combination** of finitely many elements $a_1, ..., a_n$ of M is an element of M of the form $r_1m + 1 + \cdots + r_nm_n$ for some $r_1, \ldots, r_n \in R$.

Definition 1.52. Let R be a ring with $1 \neq 0$ and let M be an R-module. For a subset A of M, the submodule of M **generated by** A is

$$RA := \{r_1a_1 + \cdots + r_na_n \mid n \geq 0, r_i \in R, a_i \in A\}.$$

We M is **generated by** A if $M = RA$. If M is an F-vector space, we say that M is **spanned by** a set A instead of generated by A.

A module M is **finitely generated** if there is a finite subset A of M that generates M. If $A = a$ has a single element, the module $RA = Ra$ is called *cyclic*.

Exercise 9. Let M be an R-module and let $A \subseteq M$. Then RA is the smallest submodule of M containing A, that is

$$RA \quad = \bigcap_{A \subseteq N, N \text{ submodule of } M} N.$$

Exercise 10. Being finitely generated and being cyclic are R-module isomorphism invariants.

Example 1.53. Let R be a ring with $1 \neq 0$.

(1) $R = R1$ is cyclic.

(2) $R \oplus R$ is generated by $\{(1, 0), (0, 1)\}$.

(3) $R[x]$ is generated as an R-module by the set $\{1, x, x^2, \ldots, x^n, \ldots\}$ of monic monomials in the variable x.

(4) Let $M = \mathbb{Z}[x, y]$. M is generated by

- $\{1, x, y\}$ as a ring,
- $\{1, y, y^2, \ldots, y^n, \ldots\}$ as an $\mathbb{Z}[x]$-module, and
- $\{x^iy^j \mid i, j \in \mathbb{Z}_{\geqslant 0}\}$ as a group ($\mathbb{Z}$-module).

Lemma 1.54. *Let R be a ring with $1 \neq 0$, let M be an R-module, and let N be an R-submodule of M.*

(1) If M is finitely generated as an R-module, then so is M/N.

(2) If N and M/N are finitely generated as R-modules, then so is M.

Proof. The proof of (2) will be a problem set question. To show (1), note that if $M = RA$ then $M/N = R\bar{A}$, where $\bar{A} = \{a + N \mid a \in A\}$. □

Definition 1.55. Let M be an R-module and let A be a subset of M. The set A is **linearly independent** if whenever $r_1, \ldots, r_n \in R$ and $a_1, \ldots, a_n$ are distinct elements of A satisfying $r_1a_1 + \cdots + r_na_n = 0$, then $r_1 = \cdots = r_n = 0$. Otherwise A is **linearly dependent**.

Definition 1.56. A subset A of an R-module M is a **basis** of M if A is linearly independent and generates M. An R-module M is a **free** R-module if M has a basis.

We will later see that over a field, every module is free. However, when R is not a field, there are R-modules that are not free; in fact, *most* modules are not free.

Example 1.57. Here are some examples of free modules:

(1) If we think of R as a module over itself, it is free with basis $\{1\}$.

(2) The module $R \oplus R$ is free with basis $\{(1,0),(0,1)\}$.

(3) The R-module $R[x]$ is free, and $\{1, x, x^2, \ldots, x^n, \ldots\}$ is a basis.

(4) Let $M = \mathbb{Z}[x,y]$. Then $\{1, y, y^2, \ldots, y^n, \ldots\}$ is a basis for the $\mathbb{Z}[x]$-module M, and $\{x^iy^j \mid i,j \in \mathbb{Z}_{\geqslant 0}\}$ is a basis for the $\mathbb{Z}$-module M.

Example 1.58. $\mathbb{Z}/2$ is not a free $\mathbb{Z}$-module. Indeed suppose that A is a basis for $\mathbb{Z}/2$ and $a \in A$. Then $2a = 0$ so A cannot be linearly independent, a contradiction.

Lemma 1.59. *If A is a basis of M then every nonzero element $0 \neq m \in M$ can be written uniquely as $m = r_1a_1 + \cdots + r_na_n$ with a_i distinct elements of A and $r_i \neq 0$.*

Proof. Suppose that if $m \neq 0$ and A_1, A_2 are finite subsets of A such that

$$m = \sum_{a \in A_1} r_a a = \sum_{b \in A_2} s_b b$$

for some $r_a, s_b \in R$. Then

$$\sum_{a \in A_1 \cap A_2} (r_a - s_a)a + \sum_{a \in A_1 \setminus A_2} r_a a - \sum_{a \in A_2 \setminus A_1} s_a a = 0.$$

Since A is a linearly independent set, we conclude that $r_a = s_a$ for $a \in A_1 \cap A_2$, $r_a = 0_R$ for $a \in A_1 \setminus A_2$, and $s_a = 0_R$ for $a \in A_2 \setminus A_1$. Set

$$B := \{a \in A_1 \cap A_2 \mid r_a \neq 0_R\}.$$

Then

$$m = \sum_{a \in B} r_a a$$

is the unique way of writing m as a linear combination of elements of A with nonzero coefficients. □

Theorem 1.60. *Let R be a ring, M be a free R-module with basis B, N be any R-module, and let $j : B \to N$ be any function. Then there is a unique R-module homomorphism $h : M \to N$ such that $h(b) = j(b)$ for all $b \in B$.*

Proof. We have two things to prove: existence and uniqueness.
Existence: By Lemma 1.59, any $0 \neq m \in M$ can be written uniquely as

$$m = r_1 b_1 + \cdots + r_n b_n$$

with $b_i \in B$ distinct and $0 \neq r_i \in R$. Define $h \colon M \to N$ by

$$\begin{cases} h(r_1 b_1 + \cdots + r_n b_n) = r_1 j(b_1) + \cdots + r_n j(b_n) & \text{if } r_1 b_1 + \cdots + r_n b_n \neq 0 \\ h(0_M) = 0_N \end{cases}$$

One can check that this satisfies the conditions to be an R-module homomorphism (exercise!).

Uniqueness: Let $h : M \to N$ be an R-module homomorphism such that $h(b_i) = j(b_i)$. Then in particular $h \colon (M, +) \to (N, +)$ is a group homomorphism and therefore $h(0_m) = 0_N$ by properties of group homomorphisms. Furthermore, if $m = r_1 b_1 + \cdots + r_n b_n$ then

$$h(m) = h(r_1 b_1 + \cdots + r_n b_n) = r_1 h(b_1) + \cdots + r_n h(b_n) = r_1 j(b_1) + \cdots + r_n j(b_n)$$

by the definition of homomorphism, and because $h(b_i) = j(b_i)$. □

Corollary 1.61. *If A and B are sets of the same cardinality, and fix a bijection $j : A \to B$. If M and N are free R-modules with bases A and B respectively, then there is an isomorphism of R-modules $M \cong N$.*

Proof. Let $g : M \to N$ and $h : N \to M$ be the module homomorphisms induced by the bijection $j : A \to B$ and its inverse $j^{-1} : B \to A$, which exist by Theorem 1.60. We will show that h and g are inverse homomorphisms. First, note that $g \circ h : N \to N$ is an R-module homomorphism and $(g \circ h)(b) = g(j^{-1}(b)) = j(j^{-1}(b)) = b$ for every $b \in B$. Since the identity map id_N is an R-module homomorphism and $id_N(b) = b$ for every $b \in B$, by the uniqueness in Theorem 1.60 we have $g \circ h = \mathrm{id}_n$. Similarly, one shows that $h \circ g = \mathrm{id}_M$. □

The corollary gives that, up to isomorphism, there is only one free module with basis A, provided such a module exists. But does a free module generated by a given set A exist? It turns out it does.

Definition 1.62. Let R be a ring and let A be a set. The free R-module generated by A, denoted $F_R(A)$ is the set of formal sums

$$\begin{aligned} F_R(A) &= \{r_1 a_1 + \cdots + r_n a_n \mid n \geqslant 0, r_i \in R, a_i \in A\} \\ &= \left\{ \sum_{a \in A} r_a a \mid r_a \in R, r_a = 0 \text{ for all but finitely many } a \right\}, \end{aligned}$$

with addition defined by

$$\left(\sum_{a \in A} r_a a \right) + \left(\sum_{a \in A} s_a a \right) = \sum_{a \in A} (r_a + s_a) a$$

and R-action defined by

$$r \left(\sum_{a \in A} r_a a \right) = \sum_{a \in A} (r r_a) a.$$

Exercise 11. This construction $F_R(A)$ results in an R-module, which is free with basis A, and $F_R(A) \cong \bigoplus_{a \in A} R$.

Theorem 1.63 (Uniqueness of rank over commutative rings)**.** *Let R be a commutative ring with $1 \neq 0$ and let M be a free R-module. If A and B are both bases for M, then A and B have the same cardinality, meaning that there exists a bijection $A \to B$.*

Proof. You will show this in the next problem set (at least in the case where M has a finite basis). □

Definition 1.64. Let R be a commutative ring with $1 \neq 0$ and let M be a free R-module. The **rank** of M is the cardinality of any basis of M.

Example 1.65. Let R be a commutative ring with $1 \neq 0$. The rank of R^n is n. Note that by Corollary 1.61, any free R-module of rank n must be isomorphic to R^n.

Earlier, we described the R-module structure on the direct sum of R-modules; this is how we construct R^n, by taking the direct sum of n copies of the R-module R. This construction can also be described as the direct product of n copies of R. However, the direct sum and direct product are two different constructions.

Definition 1.66. Let R be a ring. Let $\{M_a\}_{a \in J}$ be a collection of R-modules. The **direct product** of the R-modules M_a is the Cartesian product

$$\prod_{a \in J} M_a := \{(m_a)_{a \in J} \mid m_a \in M_a\}$$

with addition defined by

$$(m_a)_{a \in J} + (n_a)_{a \in J} := (m_a + n_a)_{a \in J}$$

and R-action defined by

$$r(m_a)_{a \in J} = (rm_a)_{a \in J}.$$

The **direct sum** of the R-modules M_a is the R-submodule $\bigoplus_{a \in J} M_a$ of the direct product $\prod_{a \in J} M_a$ given by

$$\bigoplus_{a \in J} M_a = \{(m_a)_{a \in J} \mid m_a = 0 \text{ for all but finitely many } a\}.$$

Exercise 12. The direct sum and the direct product of an arbitrary family of R-modules are R-modules.

Example 1.67. Suppose that $|A| = n < \infty$. Let $M_1, \ldots, M_n$ be R-modules. The direct product module $M_1 \times \cdots \times M_n$ is the abelian group $M_1 \times \cdots \times M_n$ with ring action given by $r(m_1, \ldots, m_n) = (rm_1, \ldots, rm_n)$ for all $r \in R$ and $m_i \in M_i$. Comparing the definitions we see that

$$M_1 \times \cdots \times M_n = M_1 \oplus \cdots \oplus M_n.$$

If $M_i = R$ for $1 \leqslant i \leqslant n$, then we denote $R^n = \underbrace{R \times \cdots \times R}_{n} = \underbrace{R \oplus \cdots \oplus R}_{n}$.

It is useful to talk about maps from the factors/summands to the direct product/ direct sum and conversely.

Definition 1.68. For $i \in J$ the *inclusion of the i-th factor* into a direct product or direct sum is the map

$$\iota_i \colon M_i \to \prod_{a \in J} M_a \text{ or } \iota_i \colon M_i \to \bigoplus_{a \in J} M_a, \iota_i(m) = (m_a)_{a \in J}, \text{ where } m_a = \begin{cases} m & \text{if } a = i \\ 0 & \text{if } a \neq i \end{cases}.$$

For $i \in J$ the i-th *projection map* from a direct product or a direct sum module is

$$\pi_i \colon \prod_{a \in J} M_a \to M_i \text{ or } \pi_i : \bigoplus_{a \in J} M_a \to M_i, \pi_i \left((m_a)_{a \in J}\right) = m_i.$$

Lemma 1.69. *Projections from direct products or sums of R-module, inclusions into direct products or sums of R-modules, and products of R-module homomorphisms are R-module homomorphisms. Furthermore, inclusions are injective, projections are surjective, and*

$$\pi_i \circ \iota_i = \mathrm{id}_{M_i}.$$

Also, $\iota_i(M_i)$ is an R-submodule of the direct product/sum which is isomorphic to M_i.

Note, however, that $\iota_i \circ \pi_i \neq \mathrm{id}$.

Chapter 2

Vector spaces and linear transformations

2.1 Classification of vector spaces and dimension

Recall that for a subset A of an F-vector space V, the **span** of A, denoted $\text{span}(A)$, is the subspace generated by A:

$$\text{span}(A) := \{\sum_{i=1}^{n} c_i a_i \mid n \geqslant 0, c_i \in F, a_i \in A\}.$$

Lemma 2.1. *Suppose I is a linearly independent subset of an F-vector space V and $v \in V \setminus \text{span}(I)$, then $I \cup \{v\}$ is also linearly independent.*

Proof. Let $w_1, \ldots, w_n$ be any list of distinct elements of $I \cup \{v\}$ and suppose that $\sum_i c_i w_i = 0$ for some $c_i \in F$. If none of the w_i's is equal to v, then $c_i = 0$ for all i, since I is linearly independent. Without loss of generality, say $w_1 = v$. If $c_1 = 0$ then $c_i = 0$ for all i by the same reasoning as in the previous case. If $c_1 \neq 0$, then

$$v = \sum_{i \geqslant 2} \frac{c_i}{c_1} w_i \in \text{span}(I),$$

contrary to assumption. This proves that $I \cup \{v\}$ is a linearly independent set. □

To prove that every vector space has a basis, we will need to use Zorn's Lemma. Before we recall what Zorn's Lemma says, let's recall some notation:

Definition 2.2. A **poset** is a set S with an order relation $\leqslant$ such that for all elements $x, y, z \in S$ we have

- $x \leqslant x$,
- if $x \leqslant y$ and $y \leqslant z$ then $x \leqslant z$, and
- if $x \leqslant y$ and $y \leqslant x$ then $x = y$.

A **totally ordered** set is a poset $(T, \leq)$ such that for all $x, y \in T$ either $x \leqslant y$ or $y \leqslant x$.

Example 2.3. Given a set X, the collection $\mathcal{P}(X)$ of all subsets of X forms a poset with $\leqslant$ defined to be set containment $\subseteq$. Unleess X is empty or a singleton, the poset $\mathcal{P}(X)$ is not totally ordered.

Definition 2.4. Let $(\mathcal{A}, \leqslant)$ be a **poset**, meaning that $\mathcal{A}$ is a set with a partial order $\leqslant$. A subset $\mathcal{B}$ of $\mathcal{A}$ is **totally ordered** if for all $b, b' \in \mathcal{B}$ either $b \leqslant b'$ or $b' \leqslant b$; a totally ordered subset of $\mathcal{A}$ is sometimes called a **chain**. We say a subset $\mathcal{B}$ of $\mathcal{A}$ has an **upper bound** in $\mathcal{A}$ if there exists an element $u_B \in \mathcal{A}$ such that $b \leqslant u_{\mathcal{B}}$ for all $b \in \mathcal{B}$. We say $\mathcal{A}$ has a **maximal element** if there exists $m \in \mathcal{A}$ such that whenever $x \in \mathcal{A}$ and $m \leqslant x$ then $m = x$.

Axiom 2.5 (Zorn's Lemma)**.** If $\mathcal{A}$ is a nonempty poset such that every totally ordered subset $\mathcal{B} \subseteq \mathcal{A}$ has an upper bound in $\mathcal{A}$, then there is a maximal element $m \in \mathcal{A}$.

Some mathematicians refuse to accept Zorn's Lemma into their axiom system. We will at least pretend to be mathematicians who do. Fun fact: Theorem 2.6 is actually equivalent to the Axiom of Choice, meaning that if one replaces the Axiom of Choice in the ZFC axioms for set theory by Theorem 2.6, that does not change set theory – and one would then be able to deduce the Axiom of Choice.

If we accept Zorn's Lemma, we can now show that every vector space has a basis.

Theorem 2.6 (Every vector space has a basis)**.** *Let V be an F-vector space and assume $I \subseteq S \subseteq V$ are subsets such that I is linearly independent and S spans V. Then there is a subset B with $I \subseteq B \subseteq S$ such that B is a basis.*

Before we prove this theorem, note that a corollary of Theorem 2.6 is that every vector space has a basis; in particular, this says that every module over a field is free!

Corollary 2.7. *Every vector space V has a basis. Moreover, every linearly independent subset of V is contained in some basis, and every set of vectors that spans V contains some basis.*

Proof. For this first part, apply the theorem with $I = \emptyset$ and $S = V$. For the second and third, use I arbitrary and $S = V$ and $I = \emptyset$ and S arbitrary, respectively. □

Example 2.8. $\mathbb{R}$ has a basis as a $\mathbb{Q}$-vector space; just don't ask me what it looks like.

We will not prove Theorem 2.6. But before we give a formal proof, let's first give a heuristic proof. To so that, start with I. If $\text{span}(I) = V$, then $B = I$ does the job. If not, then since $\text{span}(S) = V$, there must be a $v \in S \setminus \text{span}(I)$. Let $I' := I \cup \{v\}$. Then $I' \subseteq S$ and, by Lemma 2.1, I' is linearly independent. If $\text{span}(I') = V$, we have found our B, and if not we construct I'' from I' just as we constructed I' from I. At this point we would like to say that this process cannot go on for ever, and this is more-or-less true. But at least in an infinite dimensional setting, we need to use Zorn's Lemma to complete the proof rigorously.

Proof of Theorem 2.6. Let $\mathcal{P}$ denote the collection of all subsets X of V such that $I \subseteq X \subseteq S$ and X is linearly independent. We make $\mathcal{P}$ into a poset by the order relation given by set containment $\subseteq$. We note that $\mathcal{P}$ is not empty since, for example $I \in \mathcal{P}$.

Let $\mathcal{T}$ be any nonempty chain in $\mathcal{P}$. Let $Z = \bigcup_{Y \in \mathcal{T}} Y$. We claim $Z \in \mathcal{P}$. Given $z_1, \ldots, z_m \in Z$, for each i we have $z_i \in Y_i$ for some $Y_i \in T$. Since T is totally ordered, one of $Y_1, \ldots, Y_m$ contains all the others and hence contains all the z_i's. Since Y_i is linearly independent, this shows $z_1, \ldots, z_m$ are linearly independent. Thus Z is linearly independent. Since $\mathcal{T}$ is non-empty, $Z \supseteq I$ and hence $Z \in \mathcal{P}$. It is an upper bound for $\mathcal{T}$ by construction.

By Zorn's Lemma, $\mathcal{P}$ has a maximal element B, which we claim is a basis for V. Note that B is linearly independent and $I \subseteq B \subseteq S$ by construction. We need to show that it spans V. Suppose not. Since S spans V, if $S \subseteq \operatorname{span}(B)$, then $\operatorname{span}(B)$ would have to be all of V. So, there is at least one $v \in S$ such that $v \notin \operatorname{span}(B)$, and set $X := B \cup \{v\}$. Clearly, $I \subset X \subseteq S$ and, by Lemma 2.1, X is linearly independent. This shows that X is an element of $\mathcal{P}$ that is strictly bigger than B, contrary to the maximality of B. □

Corollary 2.9. *Suppose F is a field an W is a subspace of the F-vector space V. Then every basis of W extends to a basis of V, that is, if B is a basis of W then there exists a basis $\tilde{B}$ of V such that B is a subset of $\tilde{B}$.*

Proof. Apply Corollary 2.7 with $B = I$ and $S = V$. Since B is a basis of W, B is linearly independent, and B remains linearly independent when regarded as a subset of V. □

Remark 2.10. It is *not* true that, with the notation of the previous Corollary, if $\tilde{B}$ is a basis of V then there exists a basis B of W such that B is a subset of $\tilde{B}$. For instance, take $F = \mathbb{R}$, $V = \mathbb{R}^2$, $\tilde{B} = \{(1,0),(0,1)\}$ and W the subspace spanned by $(1,1)$.

Definition 2.11. A vector space is **finite dimensional** if there is spanned by a finite subset.

Thanks to Theorem 2.6, this is equivalent to the property that it has a finite basis. In the language of modules, a finite dimensional vector space is just a finitely generated F-module.

The following is an essential property of vector spaces that eventually will allow us to compare bases in terms of size.

Lemma 2.12 (Exchange Property)**.** *Let B be a basis for the vector space V and consider any finite set of linearly independent vectors $C = \{c_1, \ldots, c_m\}$ in V. Then there are distinct vectors $b_1, \ldots, b_m$ in B such that $(B \setminus \{b_1, \ldots, b_m\}) \cup C$ is also a basis V.*

Proof. Using induction on k, we will show that for each k with $0 \leqslant k \leqslant m$ there are distinct vectors $b_1, \ldots, b_k$ in B such that $(B \setminus \{b_1, \ldots, b_k\}) \cup \{c_1, \ldots, c_k\})$ is also a basis of V. In the base case, $k = 0$, there is nothing to show. The terminal case, $k = m$, gives us the desired statement.

For the inductive step, assume $B' = (B \setminus \{b_1, \ldots, b_k\}) \cup \{c_1, \ldots, c_k\})$ is also a basis of V. Since $c_{k+1} \in V$, we can write

$$c_{k+1} = \sum_{i=1}^{n} \lambda_i b_i + \sum_{i=1}^{k} \mu_i c_i$$

for some scalars $\lambda_i, \mu_i \in F$ and some elements $b_i \in B \setminus \{b_1, \ldots, b_k\}$. Note that since C is linearly independent, at least one of the scalars λ_i is nonzero. Let i_0 be such that $\lambda_{i_0} \neq 0$, and notice that solving for b_{i_0} from the displayed equation gives that $b_{i_0} \in \operatorname{span}(B'')$ where $B'' = (B' \setminus \{b_{i_0}\}) \cup \{c_{k+1}\})$. Now we can "replace" b_{i_0} by c_k, since the previous statement implies $\operatorname{span}(B'') = \operatorname{span}(B') = V$ and moreover B'' is linearly independent since otherwise B' would be linearly dependent. □

Next, we will show that all bases of the same vector space have the same cardinality. We will only prove this under the assumption that V is finite dimensional, though it is true even if V has infinite dimension.

Theorem 2.13 (Dimension Theorem)**.** *Any two bases of the same vector space have the same cardinality.*

Proof of the finite dimensional case. Suppose V is finite dimensional. Then it has a finite basis B. Let B' be any other basis, and note that we cannot yet assume B' is necessarily finite. Let $\{c_1, \ldots, c_m\}$ be any m-element subset of B' for any m. An immediate consequence of Lemma 2.12 is that $m \leqslant |B|$, since otherwise we could not find m distinct elements of B to replace the c_i's by. Since every finite subset of B' has cardinality no larger than $|B|$, this proves that B' is finite and $|B'| \leqslant |B|$. By symmetry, we obtain $|B| \leqslant |B|'$ too, hence equality follows. □

Definition 2.14. The **dimension** of a vector space V, denoted $\dim_F(V)$ or $\dim(V)$, is the cardinality of any of its bases.

Example 2.15. $\dim_F(F^n) = |\{e_1, e_2, \ldots, e_n\}| = n$.

Theorem 2.16 (Classification of finitely generated vector spaces)**.** *Let F be a field.*

(1) Every finitely generated vector space over F is isomorphic to F^n for $n = \dim_F(V)$.

(2) For any $m, n \in \mathbb{Z}_{\geqslant 0}$, $F^m \cong F^n$ if and only if $m = n$.

Proof. To show (1), let V be a finite dimensional F-vector space. Then F has a finite spanning set S and by Theorem 2.6 there is a basis $B \subseteq S$ for V. Notice that B is necessarily finite and $V = FB$. Set $|B| = n$ and $B = \{b_1, \ldots, b_n\}$. By Theorem 1.60, there is a linear transformation $f\colon F^n \to V$ such that $f(e_i) = b_i$ as well as a linear transformation $g\colon V \to F^n$ such that $g(b_i) = e_i$. Then both $f \circ g : V \to V$ and $g \circ f : F^n \to F^n$ are linear transformation which agree with the identity map on a basis. Hence by the uniqueness part of Theorem 1.60 we have $f \circ g = \mathrm{id}_V$ and $g \circ f = \mathrm{id}_{F^n}$. Therefore, these maps are the desired isomorphisms.

To show (2), let $\varphi : F^m \cong F^n$ be a vector space isomorphism and let B be a basis of F^m. We claim that $\varphi(B)$ is a basis for F^n. Indeed, if

$$\sum_{i=1}^{m} c_i \varphi(b_i) = 0 \quad \text{then} \quad \varphi\left(\sum_{i=1}^{m} c_i b_i\right) = 0, \quad \text{so} \quad \sum_{i=1}^{m} c_i b_i = 0$$

since φ is injective. But B is linearly independent, so we must have $c_i = 0$ for all $1 \leqslant i \leqslant m$. If $v \in F^n$, then since B spans F^m we have

$$\varphi^{-1}(v) = \sum_{i=1}^{m} c_i b_i$$

for some c_i. Thus

$$v = \sum_{i=1}^{m} c_i \varphi(b_i),$$

which shows $\varphi(B)$ spans F^n. By the Dimension Theorem, we have

$$\dim_F(F^n) = n = |\varphi(B)| = |B| = m.$$ □

Remark 2.17.

(1) The same proof as in part (1) of Theorem 2.16 above shows that every finitely generated free R-module is isomorphic to R^n for some $n \geqslant 0$.

(2) Part (2) of the Classification Theorem can be extended to modules over commutative rings as stated in Theorem 1.63; this is a problem in Problem Set 3.

(3) The Classification Theorem yields that dimension is an isomorphism invariant.

Corollary 2.18. *Two finite dimensional vector spaces V and V' over the same field F are isomorphic if and only if* $\dim_F(V) = \dim_F(V')$.

Proof. By Theorem 2.16, V and V' are both of the form $V \cong F^m$ and $V' \cong F^n$, while $F^m \cong F^n$ if and only if $m = n$. □

A word on infinite-dimensional vector spaces.

Example 2.19. Consider the vector space $F[x]$. This cannot be a finite dimensional vector space. For instance, if $\{f_1, \ldots, f_n\}$ were a basis, then setting

$$M = \max_{1 \leqslant j \leqslant n} \{\deg(f_j)\}$$

we see that the element x^{M+1} is not be in the span of $\{f_1, \ldots, f_n\}$. We can find a basis for this space though. Consider the collection $B = \{1, x, x^2, \ldots\}$. This set is linearly independent and spans $F[x]$, thus it forms a basis for $F[x]$. This basis is *countable*, so $\dim_F(F[x]) = \aleph_0 = |\mathbb{N}|$.

Example 2.20. Consider the real vector space

$$V := \mathbb{R}^{\mathbb{N}} = \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \cdots .$$

This space can be identified with sequences $\{a_n\}$ of real numbers. One might be interested in a basis for this vector space. At first glance, the most obvious choice for a basis would be $E = \{e_1, e_2, \ldots\}$. It turns out that E is the basis for the direct sum $\bigoplus_{i \in \mathbb{N}} \mathbb{R}$. However, it is immediate that this set does not span V, as $v = (1, 1, \ldots)$ can not be represented as a finite linear combination of these elements. Since v is not in $\text{span}(E)$, then by Lemma 2.1 we know that $E \cup \{v\}$ is a linearly independent set. However, this new set $E \cup \{v\}$ does not span BV either, as $(1, 2, 3, 4, \ldots)$ is not in the span of $E \cup \{v\}$. We know that V has a basis, but it can be shown that no countable collection of vectors forms a basis for this space, and in fact $\dim_{\mathbb{R}}(\mathbb{R}^{\mathbb{N}}) = |\mathbb{R}|$.

We now deduce some formulas that relate the dimensions of various vector spaces.

Theorem 2.21. *Let W be a subspace of a vector space V. Then*

$$\dim(V) = \dim(W) + \dim(V/W).$$

Here the dimension of a vector space is understood to be either a nonnegative integer or ∞, and the arithmetic of the formula is understood to follow the rules $n + \infty = \infty = \infty + \infty$ for any $n \in \mathbb{Z}_{\geqslant 0}$. We leave the proof for Problem Set 4.

Example 2.22. Consider the vector space $V = \mathbb{R}^2$ and its subspace $W = \text{span}\{e_1\}$. Then the quotient vector space V/W is, by definition,

$$V/W = \{(x, y) + W \mid (x, y) \in \mathbb{R}^2\}.$$

Looking at each coset we see that

$$(x, y) + W = (x, y) + \text{span}\{e_1\} = \{(x, y) + (a, 0) \mid a \in \mathbb{R}\} = \{(t, y) \mid t \in \mathbb{R}\},$$

so $(x, y) + W$ is geometrically a line parallel to the x-axis and having the y-intercept y. It is intuitively natural to identify such a line with its intercept, which gives a map

$$V/W \to \text{span}\{e_2\} \quad (x, y) + W \mapsto (0, y).$$

It turns out that this map is a vector space isomorphism, hence

$$\dim(V/W) = \dim(\text{span}\{e_2\}) = 1$$

and we can check that

$$\dim(W) + \dim(V/W) = 1 + 1 = 2 = \dim(V).$$

If V and W are both infinite dimensional vector spaces, it can happen that V/W is finite dimensional but also that it is infinite dimensional.

Example 2.23. Let $V = F[x]$, which we saw in Example 2.19 is an infinite dimensional vector space over F. Fix a polynomial f with $\deg(f) = d$, and note that the ideal (f) of $F[x]$ generated by f is also an F-vector subspace of $F[x]$ via restriction of scalars. We will show later that $\dim(F[x]/(f)) = d$. In contrast, the subspace E of all even degree polynomials in $F[x]$ together with the zero polynomial, then $\dim(F[x]/E) = \infty$.

Definition 2.24. Let $T \colon V \to W$ be a linear transformation. The **nullspace** of T is $\ker(T)$. The **rank** of T is $\dim(\text{im}(T))$.

Corollary 2.25 (Rank-Nullity Theorem)**.** *Let $f \colon V \to W$ be a linear transformation. Then*

$$\dim(\ker(f)) + \dim(\text{im}(f)) = \dim(V).$$

Proof. By the First Isomorphism Theorem for modules we have $V/\ker(f) \cong \text{im}(f)$, thus

$$\dim\left(V/\ker(f)\right) = \dim(\text{im}(V)).$$

By Theorem 2.21, we have

$$\dim(V) = \dim(\ker(V)) + \dim\left(V/\ker(f)\right).$$

Thus

$$\dim(V) = \dim(\ker(V)) + \dim\left(V/\ker(f)\right) = \dim(\ker(V)) + \dim(\text{im}(V)).$$

□

2.2 Linear transformations and homomorphisms between free modules

Exercise 13. If W is a free R-module with basis $C = \{c_1, \ldots, c_m\}$ and $w \in W$, then w can be written *uniquely* as $w = \sum_{j=1}^{m} a_j c_j$ with $a_1, \ldots, a_m \in R$.

Definition 2.26 (The matrix of a homomorphism between free modules)**.** Let R be a commutative ring with $1 \neq 0$. Let V be a finitely generated free R-modules of rank n, and let W be a finitely generated free R-module of rank m. Let $B = \{b_1, \ldots, b_n\}$ and $C = \{c_1, \ldots, c_m\}$ be *ordered* bases of V, W. Given an R-module homomorphism $f : V \to W$, we define elements $a_{ij} \in R$ for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$ by the formulas

$$f(b_i) = \sum_{j=1}^{m} a_{j,i} c_j. \tag{2.2.1}$$

The matrix

$$[f]_B^C = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

is said to **represent** the homomorphism f with respect to the bases B and C.

Remark 2.27. By Exercise 13, the coefficients $a_{j,i}$ in equation 2.2.1 are uniquely determined by the $f(b_i)$ and the elements of C. The coefficients $a_{j,i}$ corresponding to $f(b_i)$ form the ith column of $[f]_B^C$. Note that $[f]_B^C$ is an $m \times n$ matrix with entries in R.

Definition 2.28. Let V and W be finite F-vector spaces of dimension n and m with ordered bases B and C respectively and let $f : V \to W$ be a linear transformation. The matrix $[f]_B^C$ is called the **matrix of the linear transformation** f with respect to the bases B and C.

Example 2.29. If $\mathrm{id}_V : V \to V$ is the identity automorphism of an n-dimensional free R-module V, then for any basis B of V we have $\mathrm{id}_V(b_i) = b_i$ for all i and hence

$$[\mathrm{id}_V]_B^B = I_n.$$

Example 2.30. Let P_3 denote the the F-vector space of polynomials of degree at most 3 (including the zero polynomial) and consider the linear transformation $d : P_3 \to P_3$ given by taking the derivative $d(f) = f'$. Let $B = \{1, x, x^2, x^3\}$. Then

$$[f]_B^B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Example 2.31. Let F be a field and consider a linear transformation $f : V \to W$, where $V = F^n$ and $W = F^m$. Consider also the standard ordered bases B and C, i.e. $b_i = e_i \in V$ and $c_i = e_i \in W$. Then for any

$$v = \begin{bmatrix} l_1 \\ \vdots \\ l_n \end{bmatrix} = \sum_i l_i b_i$$

in V we have

$$f\left(\sum l_i b_i\right) = \sum_i l_i f(b_i).$$

Each $f(b_i)$ can be written uniquely as a linear combination of the c_j's as in (2.2.1):

$$f(b_i) = \sum_j a_{j,i} c_j.$$

Then we get

$$f(v) = \sum_i l_i \left(\sum_j a_{j,i} c_j\right) = \sum_j \left(\sum_i a_{j,i} l_i\right) c_j.$$

In other words, we have

$$f(v) = \begin{bmatrix} \sum_i a_{1,i} l_i \\ \vdots \\ \sum_i a_{m,i} l_i \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \cdot \begin{bmatrix} l_1 \\ \vdots \\ l_n \end{bmatrix} = [f]_B^C \cdot v.$$

Then for any

$$v = \sum_i l_i b_i$$

in V we have

$$f(\sum l_i b_i) = \sum_i l_i f(b_i).$$

Each $f(b_i)$ is uniquely expressible as a linear combination of the c_j's, say

$$f(b_i) = \sum_j a_{j,i} c_j.$$

Then we get

$$f(v) = \sum_i l_i \left(\sum_j a_{j,i} c_j\right) = \sum_j \left(\sum_i a_{j,i} l_i\right) c_j.$$

In other words, we have

$$f(v) = [f]_B^C \cdot v$$

where

$$[f]_B^C = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

and $[f]_B^C \cdot v$ denote the usual rule for matrix multiplication.

This says that any linear transformation $f : F^n \to F^m$ is given by multiplication by a matrix, since we noticed above that $f(v) = [f]_B^C \cdot v$. The same type of statement holds for free modules over commutative rings, and we will show it below in Theorem 2.32.

Theorem 2.32. *Let R be a commutative ring with $1 \neq 0$. Let V and W be finitely generated free R-modules of ranks n and m respectively. Fixing ordered bases B for V and C for W gives an isomorphism of R-modules*

$$\mathrm{Hom}_R(V, W) \cong \mathrm{M}_{m,n}(R) \qquad f \mapsto [f]_B^C.$$

If $V = W$, so that in particular $m = n$, and $B = C$, then the above map is an R-algebra isomorphism $\mathrm{End}_R(V) \cong \mathrm{M}_n(R)$.

Proof. Let $\varphi : \mathrm{Hom}_R(V, W) \to \mathrm{M}_{m,n}(R)$ be defined by $\varphi(f) = [f]_B^C$. We need to check that φ is a homomorphism of R-modules, which translates into $[f + g]_B^C = [f]_B^C + [g]_B^C$ and $[\lambda f]_B^C = \lambda [f]_B^C$ for any $f, g \in \mathrm{Hom}_R(V, W)$ and $\lambda \in R$. Let $A = [f]_B^C$ and $A' = [g]_B^C$. Then

$$(f + g)(b_i) = f(b_i) + g(b_i) = \sum_j a_{j,i} c_j + \sum_j a'_{j,i} c_j = \sum_j (a_{j,i} + a'_{i,j}) c_j$$

gives $[f + g]_B^C = A + A'$ and

$$(\lambda f)(b_i) = \lambda \left(\sum_j a_{j,i} c_j \right) = \sum_j (\lambda a_{j,i}) c_j$$

gives $[\lambda f]_B^C = \lambda A$. We leave the proof that for $f, g \in \mathrm{End}_R(V)$ we have $[f \circ g]_B^B = [f]_B^B [g]_B^B$ as an exercise.

Finally, the argument described in Example 2.31 also works for any ring R, and it can be adapted for any two chosen basis B and C, showing that φ is a bijection. □

Corollary 2.33. *For any field F and finite F-vector spaces V and W of dimension n and m respectively,* $\dim(\mathrm{Hom}_F(V, W)) = mn$.

Proof. The isomorphism $\mathrm{Hom}_F(V, W) \cong \mathrm{M}_{m,n}(F)$ gives

$$\dim\left(\mathrm{Hom}_F(V, W)\right) = \dim\left(\mathrm{M}_{m,n}(F)\right) = mn.$$

□

2.3 Change of basis

Definition 2.34. Let V be a finitely generated free module over a commutative ring R, and let B and C be bases of V. Let id_V be the identity map on V. Then $[\mathrm{id}_V]_B^C$ is a matrix called the **change of basis matrix** from B to C.

In Theorem 2.39 we will show that $[\mathrm{id}_V]_B^C$ is invertible with inverse $\left([\mathrm{id}_V]_B^C\right)^{-1} = [\mathrm{id}_V]_C^B$.

Example 2.35. Consider the subspace $V = P_2$ of $F[x]$ of all polynomials of degree up to 2, and the bases $B = \{1, x, x^2\}$ and $C = \{1, x-2, (x-2)^2\}$ of V. We calculate the change of basis matrix. We have

$$\begin{aligned}\mathrm{id}_V(1) &= 1,\\ \mathrm{id}_V(x) &= 2\cdot 1 + 1\cdot(x-2),\\ \mathrm{id}_V(x^2) &= 4\cdot 1 + 4\cdot(x-2) + 1\cdot(x-2)^2.\end{aligned}$$

Thus, the change of basis matrix is given by $[\mathrm{id}_V]_B^C = \begin{bmatrix}1 & 2 & 4\\ 0 & 1 & 4\\ 0 & 0 & 1\end{bmatrix}$.

Lemma 2.36. *If V, W, U are finitely generated free R-modules spaces with ordered bases B, C, and D, and if $f\colon V \to W$ and $g\colon W \to U$ are R-module homomorphisms, then*

$$[g \circ f]_D^B = [g]_D^C \cdot [f]_C^B.$$

Proof. Given $v \in V$, we have

$$(f\circ g)(v) = f(g(v)) = f([g]_B^C v) = [f]_C^D([g]_B^C v) = ([f]_C^D[g]_B^C)v,$$

so $[f\circ g]_B^B = [f]_B^B[g]_B^B$. □

Definition 2.37. Let V be a finitely generated free module over a commutative ring R. Two R-module homomorphisms $f, g : V \to V$ are **similar** if there is a bijective linear transformation $h : V \to V$ such that $g = h \circ f \circ h^{-1}$. Two $n \times n$ matrices A and B with entries in R are **similar** if there is an invertible $n \times n$ matrix P such that $B = PAP^{-1}$.

Remark 2.38. For elements $A, B \in \mathrm{GL}_n(R)$, the notions of similar and conjugate are the same.

Theorem 2.39. *Let V, W be finitely generated free modules over a commutative ring R, let B and B' be bases of V, let C and C' be bases of W, and let $f : V \to W$ be a homomorphism. Then*

$$[f]_{B'}^{C'} = [\mathrm{id}_W]_C^{C'}[f]_B^C[\mathrm{id}_V]_{B'}^B \tag{2.3.1}$$

In particular, if $g\colon V \to V$ is an R-module homomorphism, then $[g]_B^B$ and $[g]_{B'}^{B'}$ are similar.

Proof. Since $f = \mathrm{id}_W \circ f \circ \mathrm{id}_V$, by Lemma 2.36 we have

$$[f]_{B'}^{C'} = [\mathrm{id}_W]_C^{C'}[f]_B^C[\mathrm{id}_V]_{B'}^B.$$

Setting $V = W$, $B = C$, $B' = C'$, and $f = \mathrm{id}_V$ in (2.3.1) we have $[\mathrm{id}_V]_{B'}^{B'} = [\mathrm{id}_V]_B^{B'}[\mathrm{id}_V]_B^B[\mathrm{id}_V]_{B'}^B$. Notice that $[\mathrm{id}_V]_B^B = [\mathrm{id}_V]_{B'}^{B'} = I$ is the identity matrix, so the previous formula says that

$$I = [\mathrm{id}_V]_B^{B'} I [\mathrm{id}_V]_{B'}^B.$$

Setting $P = [\mathrm{id}_V]_B^{B'}$, we notice that the previous identity gives $P^{-1} = [\mathrm{id}_V]_{B'}^B$.

Now set $V = W, B = C, B' = C'$ and $f = g$ in (2.3.1) to obtain

$$[g]_{B'}^{B'} = [\mathrm{id}_V]_B^{B'}[g]_B^B[\mathrm{id}_V]_{B'}^B = P[g]_B^B P^{-1}. \qquad \square$$

We now come to certain special changes of basis and their matrices:

Definition 2.40. Let R be a commutative ring with $1 \neq 0$, let M be a free R-module of finite rank n, and let $B = \{b_1, \ldots, b_n\}$ be an ordered basis for M. An **elementary basis change operation** on the basis B is one of the following three types of operations:

1. Replacing b_i by $b_i + rb_j$ for some $i \neq j$ and some $r \in R$,
2. Replacing b_i by ub_i for some i and some unit u of R,
3. Swapping the indices of b_i and b_j for some $i \neq j$.

Definition 2.41. Let R be a commutative ring with $1 \neq 0$. An **elementary row operation** on a matrix $A \in \mathrm{M}_{m,n}(R)$ is one of the following three types of operations:

1. Adding an element of R times a row of A to a different row of A.
2. Multiplying a row of A by a unit of R.
3. Interchanging two rows of A.

Definition 2.42. Let R be a commutative ring with $1 \neq 0$. An **elementary matrix** over R is an $n \times n$ matrix obtained from I_n by applying a single elementary row operation:

1. For $r \in R$ and $1 \leqslant i, j \leqslant n$ with $i \neq j$, let $E_{i,j}(r)$ be the matrix with 1s on the diagonal, r in the (i, j) position, and 0 everywhere else.
2. For $u \in R^\times$ and $1 \leqslant i \leqslant n$ let $E_i(u)$ denote the matrix with (i, i) entry u, (j, j) entry 1 for all $j \neq i$, and 0 everywhere else.
3. For $1 \leqslant i, j \leqslant n$ with $i \neq j$, let $E_{(i,j)}$ denote the matrix with 1 in the (i, j) and (j, i) positions and in the (l, l) positions for all $l \notin \{i, j\}$, and 0 in all other entries.

Remark 2.43. Let E be an $n \times n$ elementary matrix.

- E is the change of basis matrix $[\mathrm{id}_V]_{B'}^B$, where B is any basis of V and B' is the basis obtained from B by the corresponding elementary basis change operation.
- If $A \in \mathrm{M}_{n,q}(R)$, then the product matrix EA is the result of performing the corresponding elementary row operation on A.
- If $B \in \mathrm{M}_{m,n}(R)$, then the product matrix BE is the result of performing the corresponding elementary column operation on B.

Chapter 3

Finitely generated modules over PIDs

We have seen that every module over a field is free. In contrast, whenever R is a commutative ring that is not a field, we can always construct modules that are not free. We will see that, however, every module is still a quotient of a free module. Describing that quotient explicitly is to give a presentation for the module, similarly to how we gave presentations for groups. We will study the particular case of finitely generated modules over PIDs in more detail.

3.1 Every module is a quotient of a free module

Lemma 3.1. *Given any ring R with $1 \neq 0$, any direct sum of copies of R is always a free R-module.*

Proof. Suppose that $F = \bigoplus_{i \in \Lambda} R$ is a direct sum of copies of R indexed by some set Λ. The tuples

$$e_i = (a_j)_{j \in \Lambda} \quad \text{with } a_j = 0 \text{ for all } j \neq i \text{ and } a_i = 1$$

generate F, since we can write any element as

$$(c_i)_{i \in \Lambda} = \sum_{i \in \Lambda} c_i e_i.$$

Notice that by definition $c_i \neq 0$ for only finitely many i, so the sum on the right has finitely many nonzero terms. Moreover, the e_i are linearly independent, and thus they form a basis for F. □

We will show in the next chapter that every when R is a field, every R-module is free. In contrast, we will also see that if R is a commutative ring that is not a field, there always exists an R-module that is not free – in fact, given a ring R that is not a field, one can give a very concrete recipe for building nonfree modules.

However, even though not all modules are free, what is true is that every R-module can be written as a quotient of a free module, as follows. Given a module M, first take a generating set for M, say $\Gamma = \{m_i\}_{i \in \Lambda}$. Notice that a generating set always exists: for example, we can take $\Gamma = M$, though of course that is a bit of an overkill, since it's quite likely that some elements can be obtained from linear combinations of others.

Next, we construct a free module on the set Λ; more precisely, we take a free module on as many generators as generators for M that we picked. Now the map

$$\bigoplus_{i\in\Lambda} R \xrightarrow{\pi} M$$

$$(r_i) \longmapsto \sum_{i\in\Lambda} r_i m_i.$$

Notice this map actually makes sense: the tuples (r_i) have only finitely many nonzero entries, and thus $\sum_{i\in\Lambda} r_i m_i$ is a (finite) linear combination of our chosen generators. Moreover, since we chose the m_i to be generators for M, this map π is surjective. It is also easy to check that it is an R-module homomorphism: in fact, this is the R-module homomorphism we would get from Theorem 1.60 by setting $e_i \mapsto m_i$.

By the First Isomorphism Theorem,

$$M \cong \bigoplus_{i\in\Lambda} R/\ker\pi.$$

This shows the following:

Theorem 3.2. *Every R-module is a quotient of a free R-module.*

Notice that the map π we constructed above depends on a choice of generating set for M. Given the map π corresponding to the set of generators $\Gamma = \{m_i\}$, each element in $\ker(\pi)$ is a **relation** among the generators for M: the tuple (r_i) is a relation for the generators $\{m_i\}$ if

$$\sum_{i\in\Lambda} r_i m_i = 0.$$

A nonzero relation among the m_i tells us that the set $\{m_i\}$ is linearly dependent. Thus we see that

$$\pi \text{ is injective} \iff \{m_i\} \text{ is linearly independent} \iff \{m_i\} \text{ is a basis for } M.$$

In particular, the existence of such a map π that is injective is equivalent to M being free. Since π is always surjective (as long as $\{m_i\}$ forms a generating set for M, we can now rephrase this as

$$\pi \text{ is an isomorphism} \iff \{m_i\} \text{ is a basis for } M.$$

The module M is free if and only if we can find a basis for M, thus M if M is free then M is isomorphic to a direct sum of copies of R. Since we have already shown that a direct sum of copies of R is free, we conclude the following:

Theorem 3.3. *An R-module is free if and only if it is isomorphic to a direct sum of copies of R.*

3.2 Presentations for finitely generated modules over noetherian rings

Writing a given R-module M as a quotient of a free module is giving a **presentation** for M. In 817, we studied presentations for groups; these consisted of a set of generators and a set (normal subgroup) of relations among these generators. Presentations are important for modules as well. In this case, the relations are encoded by a matrix, or equivalently by a homomorphism between a pair of free modules. We study below how the change of basis techniques can be applied to unravel the structure of a module starting with its presentation.

Definition 3.4. Let R be a commutative ring with $1 \neq 0$, let $A \in \mathrm{M}_{m,n}(R)$, and let $t_A : R^n \to R^m$ be the R-module homomorphism represented by A with respect to the standard bases. Notice that this homomorphism is given by the rule $t_A(v) = Av$. The **R-module presented by** A is the R-module $R^m/\operatorname{im}(t_A)$.

The R-module M presented by $A \in \mathrm{M}_{m,n}(R)$ has m generators and n relations. Each row of A corresponds to a generator for M, while each column encodes a relation among those generators. More precisely, the relations among the m generators are themselves *generated* by the n generators of $\operatorname{im}(t_A)$, which are the images of the standard basis of R^n by t_A.

Example 3.5. The $\mathbb{Z}$-module $M = \mathbb{Z}/6$ is presented by

$$\mathbb{Z} \xrightarrow{6} \mathbb{Z},$$

since $M \cong \mathbb{Z}/\operatorname{im}(t_6) = \mathbb{Z}/(6)$. Notice here we abused notation and wrote 6 instead of the 1×1 matrix $[6]$.

Example 3.6. Let $R = k[x, y]$, where k is a field, and $I = (x, y)$. The R-module $M = R/I$ has 1 generator, $m = 1 + I$, so we can write a presentation for M of the form $F \xrightarrow{p} R$ for some free module F and some R-module homomorphism p. To find such an F, we need to ask about the relations among the generators of M. For any $a \in I$, we have the relation $am = 0$, so I is the **module of relations** for this presentation of M.

How many generators does the module of relations have? In this case, we need 2: the relations $xm = 0$ and $ym = 0$ generate *all* the relations, since for any $a \in I$, we can write $a = rx + sy$ for some $x, y \in R$, and thus $am = 0$ can be rewritten as $r(xm) + s(ym) = 0$, which is a linear combination of the two relations $xm = 0$ and $ym = 0$. Finally, we have the following presentation for M:

$$R^2 \xrightarrow{\begin{bmatrix} x & y \end{bmatrix}} R.$$

Indeed, the image of $\begin{bmatrix} x & y \end{bmatrix}$ is (x, y), and $M \cong R/(x, y)$.

Conversely, we might be given a matrix and ask about what module it represents; one thing to keep in mind is that some presentations might be inefficient, either by having more generators or more relations than necessary. We want to answer to key questions: given a presentation for a module, how to find a more efficient presentation; and how to decide if two different presentations actually give us isomorphic modules. Keeping these goals in mind, let's try a more elaborate example.

Example 3.7. Consider the matrix

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 3 & 9 & 5 \\ 1 & -2 & 7 \\ 0 & 1 & 2 \end{bmatrix}.$$

What $\mathbb{Z}$-module M is presented by A? Formally, M is the quotient module $M = \mathbb{Z}^4/\operatorname{im}(t_A)$, where $t_! : \mathbb{Z}^3 \to \mathbb{Z}^4$ is defined by $t_A(v) = Av$. Since $\mathbb{Z}^4$ is generated by its standard basis elements $\{e_1, e_2, e_3, e_4\}$, we deduce as in Lemma 1.54 that $M = \mathbb{Z}^4/\operatorname{im}(t_A)$ is generated by the cosets of the e_i. To keep the notation short, we set $m_i = e_i + \operatorname{im}(t_A)$.

Let $N = \operatorname{im}(t_A)$ and note that N is the submodule of $\mathbb{Z}^4$ generated by the columns of A:

$$N = R\left\{ \begin{bmatrix} 2 \\ 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 9 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ 7 \\ 2 \end{bmatrix} \right\} = R\{2e_1 + 3e_2 + e_3, e_1 + 9e_2 - 2e_3 + e_4, 5e_2 + 7e_3 + 2e_4\}.$$

Since N maps to 0 under the quotient map $q\colon \mathbb{Z}^4 \to M = \mathbb{Z}^4/N$, the relations of M can be written as

$$\begin{cases} 2m_1 + 3m_2 + m_3 & = 0 \\ m_1 + 9m_2 - 2m_3 + m_4 & = 0 \\ 5m_2 + 7m_3 + 2m_4 & = 0. \end{cases}$$

We can now see that this is a rather inefficient presentation, since we can clearly use the first equation to solve for for $m_3 = -2m_1 - 3m_2$. This implies that M can be generated using only m_1, m_2 and m_4, that is

$$M = R\{m_1, m_2, m_3, m_4\} = \{m_1, m_2, m_4\}.$$

This eliminates the first equation and the latter two become

$$\begin{cases} 5m_1 + 15m_2 + m_4 & = 0 \\ -14m_2 - 16m_2 + 2m_4 & = 0 \end{cases}$$

Now we can also eliminate m_4, i.e leaving just two generators m_1, m_2 that satisfy

$$-24m_1 - 46m_2 = 0.$$

Another way to do this is to look at the matrix A and use elementary row operations to "make zeros" on the 1st and 2nd columns, as follows:

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 3 & 9 & 5 \\ 1 & -2 & 7 \\ 0 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 5 & -14 \\ 0 & 15 & -16 \\ 1 & -2 & 7 \\ 0 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 0 & -24 \\ 0 & 0 & -46 \\ 1 & 0 & 13 \\ 0 & 1 & 0 \end{bmatrix}$$

Eliminating the generators m_3 and m_4 amounts to dropping the first two columns (which are the 3rd and 4th standard basis vectors) as well as the last two rows. As we will prove

soon, this shows that the $\mathbb{Z}$-module presented by A is isomorphic to the $\mathbb{Z}$-module presented by

$$B = \begin{bmatrix} -24 \\ -46 \end{bmatrix}.$$

We can go further. Set $m_1' := m_1 + 2m_2$. Then m_1' and m_2 also form a generating set of M. The relation on m_1, m_2 translates to

$$-24m_1' + 2m_2 = 0$$

given by the matrix

$$C = E_{1,2}(-2)B = \begin{bmatrix} -24 \\ 2 \end{bmatrix}.$$

Note that we have done a row operation (subtract twice row 1 from row 2) to get from B to C. Continuing in this fashion by subtracting 12 row 2 from row 1 we also form

$$D = E_{1,2}(12)C = \begin{bmatrix} 0 \\ 2 \end{bmatrix},$$

The last matrix D presents the module $M' = \mathbb{Z}^2/\operatorname{im}(t_D)$ with generators a, b, where

$$a = e_1 + \operatorname{im}(t_D), \quad b = e_2 + \operatorname{im}(t_D)$$

and relation $2a = 0$. This module M' is isomorphic to our original module M. As we will see, this proves $M \cong \mathbb{Z} \oplus \mathbb{Z}/2$. An explicit isomorphism between M' and $\mathbb{Z} \oplus \mathbb{Z}/2$ is given by sending $\mathbb{Z}^2 \to \mathbb{Z} \oplus \mathbb{Z}/2$ by the unique $\mathbb{Z}$-module homomorphism defined by

$$e_1 \mapsto (1, 0) \text{ and } e_2 \mapsto (0, [1]_2).$$

Now notice that the kernel of this homomorphism is the submodule $(2e_2)\mathbb{Z} = \operatorname{im}(t_D)$. Then the first isomorphism theorem gives $M' = \mathbb{Z}^2/\operatorname{im}(t_D) \cong \mathbb{Z} \oplus \mathbb{Z}/2$.

Lemma 3.8. *Let R be a commutative ring with $1 \neq 0$, $A \in \mathrm{M}_{m,n}(R)$ and $B \in M_{m',n'}(R)$ for some $m, n, m', n' \geqslant 1$. Then A and B present isomorphic R-modules if B can be obtained from A by any finite sequence of operations of the following form:*

(a) an elementary row operation,

(b) an elementary column operation,

(c) deletion of the jth column and ith row of A if $Ae_j = e_i$, that is, if the jth column of A is the vector e_i,

(d) the reverse of 3: insertion of a row and column satisfying $Ae_j = e_i$,

(e) deletion of a column of all 0's,

(f) the reverse of 5: insertion of a column of all 0s.

Proof. It is sufficient to show that each individual operation gives an isomorphism, as the composition of isomorphisms is an isomorphism.

For operations (1) and (2), consider matrices A and A' where A' is obtained from A by the given elementary row/column operation, and set $M = R^m/\operatorname{im}(t_A)$ and $M' = R^{m'}/\operatorname{im}(t_{A'})$. We need to prove that there is an isomorphism $M \cong M'$.

In case (1), where we have an elementary row operation, let E be the corresponding elementary matrix. Since $A' = EA$, the isomorphism $E : R^n \to R^n$ maps $\operatorname{im}(A)$ bijectively onto $\operatorname{im}(A')$. Thus Q induces an isomorphism

$$M = R^m/\operatorname{im}(t_A) \xrightarrow{\cong} R^m/\operatorname{im}(t_{A'}) = M'.$$

In case (2), where we have an elementary column operation, let E be the corresponding elementary matrix. Since $A' = AE$ and since E is an isomorphism, we have

$$\operatorname{im}(t_{A'}) = \operatorname{im}(t_{AE}) = \operatorname{im}(t_A \circ t_E) = \operatorname{im}(t_A)$$

and so $m = m'$ and $M = R^m/\operatorname{im}(t_A) = R^{m'}/\operatorname{im}(t_{A'}) = M'$. In fact, note that for this one we get equality, not merely an isomorphism.

For case (3), we have $m' = m - 1$ and $n' = n - 1$. Since R^m is free, by the UMP for free modules there is a unique R-module homomorphism $p\colon R^m \to R^{m-1}$ sending

$$\begin{gathered} e_1 \mapsto e'_1, \dots, e_{i-1} \mapsto e'_{i-1} \\ e_i \mapsto 0 \\ e_{i+1} \mapsto e'_i, \dots, e_m \mapsto e'_{m-1} \end{gathered}$$

Similarly, there is a unique R-module homomorphism $q\colon R^n \to R^{n-1}$ sending

$$\begin{gathered} e_1 \mapsto e'_1, \dots, e_{j-1} \mapsto e'_{j-1}, \\ e_j \mapsto 0, \\ e_{j+1} \mapsto e'_j, \dots, e_n \mapsto e'_{n-1}. \end{gathered}$$

Here the elements e_i are part of a standard basis for R^n or for R^m, while the elements e'_i are part of a standard basis for R^{n-1} or for R^{m-1}. Then the diagram

$$\begin{array}{ccc} R^n & \xrightarrow{A} & R^m \\ {\scriptstyle q}\big\downarrow & & \big\downarrow{\scriptstyle p} \\ R^{n-1} & \xrightarrow[A']{} & R^{m-1} \end{array}$$

commutes by the definition of A'. In particular, $p(\operatorname{im}(t_A)) \subseteq \operatorname{im}(t_{A'})$ and so p induces an R-module homomorphism

$$\overline{p}\colon M \to M',$$

and we claim $\overline{p}$ is bijective.

Since p is onto, so is $\overline{p}$. Suppose $m \in \ker(\overline{p})$. Then $m = v + \operatorname{im}(t_A)$ for some $v \in R^m$ and $p(v) \in \operatorname{im}(t_{A'})$. Say $p(v) = A'w$. Since q is onto, $w = q(u)$ for some u. Then

$$p(v - Au) = p(v) - pA(u) = p(v) - A'q(u) = p(v) - A'w = p(v) - p(v) = 0,$$

and thus $v - Au \in \ker(p)$. Now, the kernel of p is clearly Re_i, so that $v - Au = re_i$ for some r. Finally, since $Ae_j = e_i$, we have $A(re_j) = re_i = v - Au$ and hence $v = A(u + re_j)$, which proves $v = t_A(u + re_j) \in \mathrm{im}(t_A)$ and hence that $m = 0$.

For (5), it is clear that the columns of A' generate the same submodule of R^m as do the columns of A, and thus $M = M'$.

Finally, for operations (4) and (6), since the isomorphism relation is reflexive, the statements of parts (3) and (5) show that parts (4) and (6) are true as well. □

Which modules have presentations? The discussion in Section 3.1 shows that the answer is every module. But if we want to make the presentation be finite (that is, so that the matrix describing the module has finitely many rows and columns) then we need to restrict ourselves to finitely generated modules. This in general does not suffice to guarantee that there will only be finitely many generators for the submodule of relations.

It might seem like no submodule of a finitely generated module could ever fail to itself be finitely generated, but indeed this happens! Before we see an example, let us introduce the class of rings over which this does *not* happen.

Definition 3.9 (Noetherian ring)**.** A ring R is **noetherian** if every ascending chain of ideals

$$I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$$

eventually stabilizes: there is some N for which $I_n = I_{n+1}$ for all $n \geqslant N$.

The following characterization of noetherian rings, which you will prove in Problem Set 5, is the key to guaranteeing that submodules of finitely generated modules are also finitely generated.

Theorem 3.10. *A commutative ring R is noetherian if and only if every ideal of R is finitely generated.*

Theorem 3.11. *Let R be a commutative ring. If R is a noetherian ring, then every submodule of a finitely generated module is also finitely generated.*

Proof. We will first prove that for each $n \geqslant 1$, every submodule of R^n is finitely generated. The base case $n = 1$ holds by Theorem 3.10, since a submodule of R^1 is the same thing as an ideal of R.

Assume $n > 1$ and that every submodule of R^{n-1} is finitely generated. Let M be any submodule of R^n. Define

$$\pi \colon R^n \twoheadrightarrow R^1$$

to be the projection onto the last component of R^n. The kernel of π_n may be identified with R^{n-1} and so $N := \ker(\pi) \cap M$ is a submodule of R^{n-1}, and it is therefore finitely generated by assumption. The image $\pi(M)$ of M under π is a submodule of R^1, that is, an ideal of R, and so it too is finitely generated by Theorem 3.10. Furthermore, by the first isomorphism theorem $M/\ker(\pi) \cong \pi(M)$ is also finitely generated. By Lemma 1.54, we deduce that M is a finitely generated module.

Now let T be any finitely generated R-module and $N \subseteq T$ any submodule. Since T is finitely generated, there exists a surjective R-module homomorphism $q \colon R^n \twoheadrightarrow T$ for some n.

Then $q^{-1}(N)$ is a submodule of R^n and hence it is finitely generated by the case we already proved, say by element $v_1, \dots, v_m \in q^{-1}(N)$. We claim that $q(v_1), \dots, q(v_m)$ generate N. Given any $a \in N$, since q is surjective we can find some $b \in q^{-1}(N)$ such that $q(b) = a$. Since $v_1, \dots, v_m$ generated $q^{-1}(N)$, we can find $c_1, \dots, c_m \in R$ such that

$$b = c_1 v_1 + \cdots + c_m v_m \implies c_1 q(v_1) + \cdots + c_m q(v_m) = q(c_1 v_1 + \cdots + c_m v_m) = q(b) = a. \quad \square$$

In fact, the converse of Theorem 3.11 is also true. More precisely, a commutative ring R is noetherian if and only if every submodule of a finitely generated module is also finitely generated.

Example 3.12. Let R be a commutative ring. Note that R is a module over itself and a submodule of R is exactly the same thing as an ideal. This module R is always finitely generated as an R-module: 1 generates R, for example. If R is not noetherian, then by Theorem 3.10 R has an ideal I that is not finitely generated. Then I is a submodule of a finitely generated module that fails to be finitely generated.

Theorem 3.13. *Any finitely generated module M over a noetherian ring R has a finite presentation given by an $m \times n$ matrix A, that is, there is an isomorphism*

$$M \cong R^m / \operatorname{im}(t_A),$$

where $t_A \colon R^n \to R^m$ is the map on free modules $t_A(v) = Av$ induced by A.

Proof. Let M be a finitely generated module over a noetherian. We start by following the general argument we described in Section 3.1: we choose a finite generating set $y_1, \dots y_m$ of M and obtain an R-module map

$$\pi \colon R^m \to M$$

that sends e_i to y_i, by using the UMP for free modules. Since every element in M is given as a linear combination of the y_i, the map π is surjective. Notice, however, that this representation as a linear combination of the y_i is not necessarily unique, so π might have a nontrivial kernel.

Since R^m is finitely generated and R is noetherian, by Theorem 3.11 the submodule $\ker(\pi)$ is also finitely generated, say by $z_1, \dots, z_n$. This too leads to a surjective R-module map $g \colon R^n \to \ker(\pi)$ that sends $e_i \mapsto z_i$. The composition of $g \colon R^n \twoheadrightarrow \ker(\pi)$ followed by the inclusion of $\iota \colon \ker(\pi) \hookrightarrow R^m$ is an R-module homomorphism $t = \iota \circ g : R^n \to R^m$ and hence by Theorem 2.32 t is given by a $m \times n$ matrix $A = [t]_B^C$ with respect to the standard bases of R^m and R^n respectively, meaning $t = t_A$.

It remains to show that $M \cong R^m / \operatorname{im}(t_A)$. First note that since $t_A = \iota \circ g$ and g is surjective we have

$$\operatorname{im}(t_A) = \operatorname{im}(\iota \circ g) = \iota(\operatorname{im}(g)) = \iota(\ker(\pi)) = \ker(\pi).$$

By the first isomorphism theorem we now have

$$M = \operatorname{im}(\pi) \cong R^m / \ker(\pi) = R^m / \operatorname{im}(t_A). \qquad \square$$

Now that we know that noetherianity is a nice condition, what rings are actually noetherian? Fortunately, the answer is many.

Example 3.14.

1. Every field k is noetherian, since (0) and k are the only ideals.
2. If R is a PID, then by definition every ideal is generated by a single element, and hence R is noetherian.
3. If R is noetherian, then you will show in Problem Set 5 that every quotient of R is also noetherian.

For more examples, the following famous theorem is useful.

Theorem 3.15 (Hilbert Basis Theorem)**.** *If R is a noetherian ring, then so is $R[x_1, \ldots, x_n]$ for all integers $n \geqslant 1$.*

In the interest of time, and since we really won't need it in this class, I will not give a proof of the Hilbert Basis Theorem.

Combining the facts above together gives the following very nice fact:

Corollary 3.16. *Let k be a field and let I be an ideal in $S = k[x_1, \ldots, x_n]$ for some $n \geqslant 1$. Then the ring S/I is noetherian.*

This includes a large collection of the rings that are of most interest in the fields of commutative algebra and algebraic geometry.

3.3 Classification of finitely generated modules over PIDs

Since any PID R is a noetherian ring, any finitely generated R-module M has a finite presentation matrix A. We will discuss a canonical form for such a matrix A and the consequences it has on determining the isomorphism type of M.

Theorem 3.17 (Smith Normal Form (SNF))**.** *Let R be a PID and let $A \in \mathrm{M}_{m,n}(R)$. Then there is a sequence of elementary row and column operations that transform A into a matrix $M = [a_{ij}]$ such that all nondiagonal entries of M are 0 and the diagonal entries of M satisfy*

$$a_{11} \mid a_{22} \mid a_{33} \mid \cdots .$$

Moreover, the number ℓ of nonzero entries of A is uniquely determined by B, and the nonzero diagonal entries $a_{11}, \ldots, a_{\ell\ell}$ are unique up to multiplication by units.

Example 3.18. If A is a 1×2 matrix, the existence portion of the Smith Normal Form amounts to the Euclidean algorithm: given (x, y), by subtracting a multiple of the one entry to the other in the usual back-and-forth way, we eventually arrive at $(\gcd(x, y), 0)$.

The general proof of Smith Normal Form is a sort of extended Euclidean algorithm.

Proof of Theorem 3.17. To prove existence, we claim there is a sequence of row and column operations that transforms A into

$$\begin{bmatrix} g & 0 \\ 0 & B \end{bmatrix}$$

for some $(n-1) \times (m-1)$ matrix B, where $g = \gcd(A)$. We adopt the convention that if A is the matrix of all zeroes, then $g = 0$. Granting this, we are done: notice that g divides every entry of B, and so by applying this fact over and over again we arrive at a matrix of the desired form M.

Let a be the upper left entry of A. Suppose a happens to be $g = \gcd(A)$. Then, in particular, it divides every entry of the first row and column of A, and so by doing row and column operations, we may zero out these entries to arrive at a matrix of the desired form directly.

We now proceed by induction on the number of prime factors of a/g. If there are no prime factors, then $a = \gcd(A)$, and we already did this case. Otherwise, there is at least one entry $b = a_{i,j}$ such that $a \nmid b$. If we can find such a b belonging to the first row of B, then we may implement the Euclidean algorithm in the form of suitable column operations, to replace a by $\gcd(a, b)$ and b by 0. Since $\gcd(a, b)/g$ has fewer prime factors than a/g, we are done by induction. Likewise if there exists such a b in the first column, we are done by induction.

The remaining possibility is that a divides every entry of the first row and first column. In this case, as before we can zero them out by row and column operations to obtain a matrix of the form

$$C = \begin{bmatrix} a & 0 \\ 0 & E \end{bmatrix}.$$

Since $\gcd(C) = \gcd(A)$, there is some element e of E such that $a \nmid e$. A suitable row operation puts e into row one without affecting a. We have thus reduced the problem to a previously solved case.

We sketch a proof of uniqueness next. Since R is a PID, the gcd of any collection of elements of R is by definition any one of the principal generators of the ideal generated by the collection.

For any i and any matrix B, let $\gcd_i(B)$ denote the gcd of all the $i \times i$ minors of B. We will not prove this, but it is true and not hard to see that,

Now here is where our proof becomes only a sketch: we will use the fact that for any i and any commutative ring R, the ideal generated by the $i \times i$ minors of a matrix with entries in R is unchanged by row and column operations. We will not prove this, but it is not hard to see. As a corollary, we obtain that for a PID, $\gcd_i$ is unchanged by row and column operations.

For a matrix of the form M, the only minors that are nonzero are those involving the same choices of columns and rows, and hence the only nonzero $i \times i$ minors of M are $g_{s_1} \cdots g_{s_i}$ for some $s_1 < \cdots < s_i$. Since $g_{s_1} \cdots g_{s_i}$ divide each other, it follows that

$$\gcd_i(A) = \gcd_i(M) = g_1 \cdots g_i.$$

In particular, the largest value of i such that some $i \times i$ minor of A is nonzero is ℓ. Also, we have

$$g_i = \frac{\gcd_i(A)}{\gcd_{i-1}(A)}.$$

This proves uniqueness, for it shows that $\ell, g_1, \ldots, g_\ell$ are all defined from A directly, without any choices. □

We now proceed to classify modules over PIDs using the Smith Normal Form for their presentation matrix. First, we need a lemma on how to interpret the module presented by a matrix in Smith Normal Form.

Lemma 3.19. *Let R be a commutative ring with $1 \neq 0$, let $m \geqslant n$, let $A = [a_{ij}] \in \mathrm{M}_{m,n}(R)$ be a matrix such that all nondiagonal entries of A are 0, and let M be the R-module presented by A. Then $M \cong R^{m-n} \oplus R/(a_{11}) \oplus \cdots \oplus R/(a_{nn})$.*

Theorem 3.20 (Classification of finitely generated modules over a PID using invariant factors)**.** *Let R be a PID and let M be a finitely generated module. Then there exist $r \geqslant 0$, $k \geqslant 0$, and nonzero nonunit elements $d_1, \ldots, d_k$ of R satisfying $d_1 \mid d_2 \mid \cdots \mid d_k$ such that*

$$M \cong R^r \oplus R/(d_1) \oplus \cdots \oplus R/(d_k).$$

Moreover r and k are uniquely determined by M, and the d_i are unique up to associates.

Proof. By Theorem 3.13, M has a presentation matrix A and by Theorem 3.17 this matrix can be put into Smith Normal Form M, where the diagonal entries of M are $d_1, \ldots, d_k$ and satisfy $d_1 \mid d_2 \mid \cdots \mid d_k$. Moreover, k is unique and the d_i are uniquely determined up to associates by A, hence by M. By Theorem 3.13, M is isomorphic to the module presented by M. By Lemma 3.19, this is isomorphic to

$$M \cong R^r \oplus R/(d_1) \oplus \cdots \oplus R/(d_k).$$

□

Definition 3.21. Let R be a PID, let $r \geqslant 0, k \geqslant 0$, and let $d_1, \ldots, d_k$ be nonzero nonunit elements of R satisfying $d_1 \mid d_2 \mid \cdots \mid d_k$. Let M be any R-module such that

$$M \cong R^r \oplus R/(d_1) \oplus \cdots \oplus R/(d_k).$$

The **free rank** of M is the integer r. The elements $d_1, \ldots, d_k$ are the invariant factors of M.

Remark 3.22. The classification theorem can be interpreted as saying that M decomposes into a free submodule R^r and a torsion submodule $\mathrm{Tor}(M) = R/(d_1) \oplus \cdots \oplus R/(d_k)$.

Example 3.23. Consider the $\mathbb{Z}$-module M presented by the matrix

$$A = \begin{bmatrix} 1 & 6 & 5 & 2 \\ 2 & 1 & -1 & 0 \\ 3 & 0 & 3 & 0 \end{bmatrix}.$$

The Smith normal form of A is

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 6 & 0 \end{bmatrix},$$

with invariant factor $d_1 = 6$. Therefore we have $M \cong \mathbb{Z}/(1) \oplus \mathbb{Z}/(1) \oplus \mathbb{Z}/(6) \cong \mathbb{Z}/(6)$.

Corollary 3.24 (Invariant factor form)**.** *Let G be a finitely generated abelian group. Then*

$$G \cong \mathbb{Z}^r \oplus (\mathbb{Z}/n_1) \oplus \cdots \oplus (\mathbb{Z}/n_k)$$

for some $r \geqslant 0$, $k \geqslant 0$, and $n_i \geqslant 2$ for all i, satisfying $n_{i+1} \mid n_i$ for all i. Moreover, the integers r, k, and $n_1, \ldots n_k$ are uniquely determined by G.

Here is a spinoff of the classification theorem.

Theorem 3.25 (Classification of finitely generated modules over a PID using elementary divisors)**.** *Let R be a PID and let M be a finitely generated module. Then there exist $r \geqslant 0$, $k \geqslant 0$, prime elements $p_1, \ldots, p_s$ of R (not necessarily distinct), and $e_1, \ldots, e_s \geqslant 1$ such that*

$$M \cong R^r \oplus R/(p_1^{e_1}) \oplus \cdots \oplus R/(p_s^{e_s}).$$

Moreover, r and k are uniquely determined by M, and the list $p_1^{e_1}, \ldots, p_s^{e_s}$ is unique up to associates and reordering.

Proof. First write M in invariant factor form $M \cong R^r \oplus R/(d_1) \oplus \cdots \oplus R/(d_k)$ then write each invariant factor as a product of prime powers $d_i = \prod_{j=n_i}^{n_{i+1}} p_j^{e_j}$ and recall that by the CRT we have

$$R/(d_i) \cong R/(p_{n_i}^{e_{n_i}}) \oplus \cdots \oplus R/(p_{n_{i+1}}^{e_{n_{i+1}}}).$$

Substituting this into the invariant factor form gives the desired result. Uniqueness follows from the uniqueness of the invariant factor form and of the prime factorizations of the d_i's. □

Definition 3.26. Let R be a PID, let $r \geqslant 0$, $s \geqslant 0$, $p_1, \ldots, p_s$ be prime elements of R, and let $e_1, \ldots, e_s \geqslant 1$. Let M be the R-module $M \cong R^r \oplus R/(p_1^{e_1}) \oplus \cdots \oplus R/(p_s^{e_s})$. The elements $p_1^{e_1}, \ldots, p_s^{e_s}$ of R are the **elementary divisors** of M.

Example 3.27. Continuing with $M \cong \mathbb{Z}/(6)$ from the previous example, we have $M \cong \mathbb{Z}/(6) \cong \mathbb{Z}/(2) \oplus \mathbb{Z}/(3)$, so the elementary divisors are 2 and 3.

Corollary 3.28. *Let G be a finitely generated abelian group. Then there exist $r, s \geqslant 0$, prime integers $p_1, \ldots, p_s$, and positive integers $e_i \geqslant 1$ such that*

$$G \cong \mathbb{Z}^r \oplus \mathbb{Z}/p_1^{e_1} \oplus \cdots \oplus \mathbb{Z}/p_s^{e_s}.$$

Moreover, r, p_i, and e_i are all uniquely determined by G.

3.4 Canonical forms for endomorphisms

Index