Math 533 Winter 2021, Lecture 9: Modules

website: `https://www.cip.ifi.lmu.de/~grinberg/t/21w/`

1. Modules ([DF, Chapter 10])

1.1. Module morphisms ([DF, §10.2]) (cont'd)

In the previous lecture, we have defined module morphisms, and discussed a few examples. Let me give one more, slightly confusing example of module morphisms. Namely, I claim that any ring morphism is a module morphism, as long as the module structures are defined correctly (warning: these are often not the module structures you expect!). To wit:

- Let R and S be two rings. Let $f : R \to S$ be a ring morphism. As we have seen in Lecture 8, the ring S then becomes a left R-module, with the action of R on S being defined by

  $$rs = f(r)s \qquad \text{for all } r \in R \text{ and } s \in S.$$

  This action is called the action on S induced by f. It is now easy to see that f is a left R-module morphism from R to S.

  Here is an example. There is a ring morphism $f : \mathbb{C} \to \mathbb{C}$ that sends each complex number $z = a + bi$ (with $a, b \in \mathbb{R}$) to its complex conjugate $\overline{z} = a - bi$. Thus, from the previous paragraph, we can conclude that this morphism f is a $\mathbb{C}$-module morphism from $\mathbb{C}$ to $\mathbb{C}$. But this is only true if the $\mathbb{C}$-module structure on the target (but not on the domain) is the one induced by f (so it is given by $rs = f(r)s = \overline{r}s$ for all $r \in \mathbb{C}$ and $s \in \mathbb{C}$), which is of course a rather nonstandard choice of a $\mathbb{C}$-module structure on $\mathbb{C}$. So f is indeed a $\mathbb{C}$-module morphism from $\mathbb{C}$ to $\mathbb{C}$, but these are two different $\mathbb{C}$-modules $\mathbb{C}$!

  Of course, writing things like this is just inviting confusion. To avoid this confusion, you need to introduce a new notation for the nonstandard $\mathbb{C}$-module $\mathbb{C}$ (the one induced by f). Namely, let us denote this new $\mathbb{C}$-module by $\overline{\mathbb{C}}$, while the unadorned symbol $\mathbb{C}$ will always mean the old, obvious $\mathbb{C}$-module structure on $\mathbb{C}$ (in which the action is just the multiplication). Thus, what we said in the previous paragraph can be restated as follows: The map f is a $\mathbb{C}$-module morphism from $\mathbb{C}$ to $\overline{\mathbb{C}}$. Actually, it is easy to see that f is a $\mathbb{C}$-module **isomorphism** from $\mathbb{C}$ to $\overline{\mathbb{C}}$. Thus, the $\mathbb{C}$-modules $\mathbb{C}$ and $\overline{\mathbb{C}}$ are isomorphic (but still should not be identified to prevent confusion).

  More generally, since $f : \mathbb{C} \to \mathbb{C}$ is a ring morphism, we can restrict any $\mathbb{C}$-module M to $\mathbb{C}$ using f. This means the following: If M is a $\mathbb{C}$-module, then we define a new $\mathbb{C}$-module structure on M by

$$rm = f(r)\, m = \overline{r}m \qquad \text{for all } r \in \mathbb{C} \text{ and } m \in M$$

(where the "rm" on the left hand side refers to the new $\mathbb{C}$-module structure, whereas the "$f(r)\, m$" and "$\overline{r}m$" refer to the old one). This new $\mathbb{C}$-module is called $\overline{M}$ (since calling it M would be asking for trouble). It is a "twisted version" of M: It is identical to M as an abelian group, but the action of $\mathbb{C}$ on it has been "twisted" (in the sense that scaling by z on $\overline{M}$ is the same as scaling by $\overline{z}$ on M).

Here is a nice thing about these twisted $\mathbb{C}$-modules: If V and W are two $\mathbb{C}$-modules (i.e., $\mathbb{C}$-vector spaces), then a $\mathbb{C}$-module morphism $f : V \to \overline{W}$ is what is known as an **antilinear map** from V to W in linear algebra. Thus, antilinear maps are "secretly" just linear maps, once you have twisted the vector space structure on the target.

We shall now state a bunch of general facts about module morphisms that are analogous to some facts we have previously stated for ring morphisms. I won't distract you with the proofs, as they are all straightforward.

As before, we fix a ring R.

Proposition 1.1.1. Let M and N be two left R-modules. Let $f : M \to N$ be an invertible left R-module morphism. Then, f is a left R-module isomorphism.

Proposition 1.1.2. Let M, N and P be three left R-modules. Let $f : N \to P$ and $g : M \to N$ be two left R-module morphisms. Then, $f \circ g : M \to P$ is a left R-module morphism.

Proposition 1.1.3. Let M, N and P be three left R-modules. Let $f : N \to P$ and $g : M \to N$ be two left R-module isomorphisms. Then, $f \circ g : M \to P$ is a left R-module isomorphism.

Proposition 1.1.4. Let M and N be two left R-modules. Let $f : M \to N$ be a left R-module isomorphism. Then, $f^{-1} : N \to M$ is a left R-module isomorphism.

Corollary 1.1.5. The relation $\cong$ for left R-modules is an equivalence relation.

Left R-module isomorphisms preserve all "intrinsic" properties of left R-modules (just like as morphisms do for properties of rings). For example, if M and N are two isomorphic left R-modules, then M has as many R-submodules as N does (and there is a one-to-one correspondence between the R-submodules of M and those of N).

All of this holds just as well for right R-modules; by now this is so obvious that we don't even need to say it. (Besides, as you have seen from exercise 2 **(d)** on homework set #2, right R-modules can be transformed into left R^{op}-modules for a certain ring R^{op}. This can also be done in reverse, and thus provides a dictionary between left modules and right modules, which can always be used to translate a statement about one kind of modules into a statement about the other. Module morphisms behave as one would expect under this dictionary: When we use this dictionary to turn two right R-modules M and N into left R^{op}-modules, the right R-module morphisms from M to N become the left R^{op}-module morphisms from M to N. This gives you all excuses you might ever need to ignore right R-modules and only work with left R-modules, until you actually need certain "hybrid" modules with both left and right structures.)

Next, we shall study kernels and images of module morphisms.

Definition 1.1.6. Let M and N be two left R-modules. Let $f : M \to N$ be a left R-module morphism. Then, the **kernel** of f (denoted $\ker f$ or $\operatorname{Ker} f$) is defined to be the subset

$$\operatorname{Ker} f := \{a \in M \mid f(a) = 0_N\}$$

of M.

Some examples:

- Let R be a commutative ring. Let $b \in R$ be nonzero. Then, the map $R \to R$, $r \mapsto br$ is an R-module morphism (check this!). The kernel of this map is
$$\{r \in R \mid br = 0\}.$$
Thus, this kernel is $\{0\}$ if and only if b is not a zero divisor.

- Both $\mathbb{Z}^3$ and $\mathbb{Z} \times (\mathbb{Z}/2)$ are $\mathbb{Z}$-modules (since we have seen in Lecture 8 that every additive group is a $\mathbb{Z}$-module). The map
$$\mathbb{Z}^3 \to \mathbb{Z} \times (\mathbb{Z}/2),$$
$$(a,b,c) \mapsto \left(a-b, \overline{b-c}\right)$$
is a $\mathbb{Z}$-module morphism. Its kernel is
$$\begin{aligned} &\left\{(a,b,c) \in \mathbb{Z}^3 \mid \left(a-b, \overline{b-c}\right) = 0_{\mathbb{Z}\times(\mathbb{Z}/2)}\right\} \\ &= \left\{(a,b,c) \in \mathbb{Z}^3 \mid a-b = 0 \text{ and } \overline{b-c} = 0\right\} \\ &= \left\{(a,b,c) \in \mathbb{Z}^3 \mid a-b = 0 \text{ and } b-c \equiv 0 \bmod 2\right\} \\ &= \left\{(a,b,c) \in \mathbb{Z}^3 \mid a = b \text{ and } b \equiv c \bmod 2\right\}. \end{aligned}$$

Kernels are a standard concept in linear algebra, where they are also called **nullspaces**. The following facts should be familiar from abstract linear algebra:

Theorem 1.1.7. Let M and N be two left R-modules. Let $f : M \to N$ be a left R-module morphism. Then, the kernel $\operatorname{Ker} f$ of f is an left R-submodule of M, whereas the image $\operatorname{Im} f = f(M)$ of f is a left R-submodule of N.

Lemma 1.1.8. Let M and N be two left R-modules. Let $f : M \to N$ be a left R-module morphism. Then, f is injective if and only if $\operatorname{Ker} f = \{0_M\}$.

We can now define quotient modules of left R-modules, in more or less the same way as we defined quotient rings of rings (but this time we need to establish an action instead of a multiplication on the quotient):

Definition 1.1.9. Let I be a left R-submodule of a left R-module M. Thus, I is a subgroup of the additive group $(M, +, 0)$, hence a normal subgroup (since $(M, +, 0)$ is abelian). Therefore, the quotient group M/I itself becomes an abelian group. Its elements are the cosets $r + I$ of I in M.

Note that the addition on M/I is given by

$$(a + I) + (b + I) = (a + b) + I \qquad \text{for all } a, b \in M.$$

We now define an action of R on M/I by setting

$$r(a + I) = ra + I \qquad \text{for all } r \in R \text{ and } a \in M.$$

(See below for a proof that this is well-defined.)

The set M/I, equipped with the addition and the action we just defined and with the element $0 + I$ as zero vector, is a left R-module. This left R-module is called the **quotient left R-module** of M by the submodule I; it is also pronounced "M **modulo** I". It is denoted M/I (so when you hear "the left R-module M/I", it always means the set M/I equipped with the structure just mentioned).

The cosets $r + I$ are called **residue classes** modulo I, and are often denoted $r \bmod I$ or $[r]_I$ or $[r]$ or $\overline{r}$. (The last two notations are used when I is clear from the context.)

Theorem 1.1.10. Let M and I be as in Definition 1.1.9. Then, the action of R on M/I is well-defined, and M/I does indeed become a left R-module when endowed with the operations and elements just described.

Theorem 1.1.11. Let I be a left R-submodule of a left R-module M. Consider the map

$$\pi : M \to M/I, \qquad a \mapsto a + I.$$

Then, π is a surjective module morphism with kernel I. This morphism π is called the **canonical projection** from M onto M/I.

Theorem 1.1.12 (Universal property of quotient modules)**.** Let M be a left R-module. Let I be a left R-submodule of M.

Let N be a left R-module. Let $f : M \to N$ be a left R-module morphism. Assume that $f(I) = 0$ (this is shorthand for saying that $f(a) = 0$ for all $a \in I$). Consider the canonical projection $\pi : M \to M/I$. Then, there is a unique left R-module morphism $f' : M/I \to N$ satisfying $f = f' \circ \pi$.

Just to unravel the abstract definition: This morphism f' sends each coset (= residue class) $a + I \in M/I$ to $f(a)$.

Theorem 1.1.13 (First isomorphism theorem for modules)**.** Let M and N be two left R-modules. Let $f : M \to N$ be a left R-module morphism. Recall that $\operatorname{Ker} f$ is an R-submodule of M, and that $\operatorname{Im} f = f(M)$ is an R-submodule of N. We have

$$M/\operatorname{Ker} f \cong f(M).$$

More precisely, the universal property of quotient modules (applied to $I = \operatorname{Ker} f$) yields a left R-module morphism $f' : M/\operatorname{Ker} f \to N$, which (if we restrict its target to its actual image $f(M)$) is a left R-module isomorphism from $M/\operatorname{Ker} f$ to $f(M)$.

Just to unravel this abstract definition: This isomorphism sends each coset (= residue class) $a + \operatorname{Ker} f \in M/\operatorname{Ker} f$ to $f(a)$. So you can reword the conclusion of Theorem 1.1.13 as follows: The map

$$\begin{aligned} M/\operatorname{Ker} f &\to f(M), \\ a + \operatorname{Ker} f &\mapsto f(a) \end{aligned}$$

is well-defined and is a left R-module isomorphism.

All results we have stated so far about modules are analogues of known results about rings. So are their proofs (which is why we are omitting them all). The Second and the Third isomorphism theorem for rings (which you have seen on homework set #2) also have analogues for modules.

Remark 1.1.14. If you have done some abstract linear algebra, the formula $M/\operatorname{Ker} f \cong f(M)$ in Theorem 1.1.13 might remind you of something.

Indeed, let R be a field. Thus, R-modules are R-vector spaces. Let M and N be two finite-dimensional R-vector spaces. Let $f : M \to N$ be a linear map. Thus, Theorem 1.1.13 yields that $M/\operatorname{Ker} f \cong f(M)$ as R-modules (i.e., as R-vector spaces). However, isomorphic vector spaces have equal dimension. Hence, from $M/\operatorname{Ker} f \cong f(M)$, we obtain

$$\dim(M/\operatorname{Ker} f) = \dim(f(M)). \tag{1}$$

However, it is not hard to see (we will see it soon) that $\dim(M/I) = \dim M - \dim I$ whenever I is a vector subspace of M. (The idea behind this formula is that when you pass from M to M/I, you are "collapsing" the "dimensions" contained in I (since you are equating any vector in I with 0), and thus the dimension of the vector space should go down by $\dim I$. Formally speaking, this can be shown using bases. We will do so below.)

As a consequence of the $\dim(M/I) = \dim M - \dim I$ formula, we have

$$\dim(M/\operatorname{Ker} f) = \dim M - \dim(\operatorname{Ker} f).$$

Hence,

$$\dim M - \dim(\operatorname{Ker} f) = \dim(M/\operatorname{Ker} f) = \dim(f(M)) \qquad \text{(by (1))}.$$

This is the **rank-nullity formula** from linear algebra (indeed, $\dim(\operatorname{Ker} f)$ is called the **nullity** of f, whereas $\dim(f(M))$ is called the **rank** of f).

1.2. Spanning, linear independence, bases and free modules ([DF, §10.3])

We shall now generalize some classical notions from linear algebra (spanning, linear independence and bases) to arbitrary R-modules.

Let us still fix a ring R.

Definition 1.2.1. Let M be a left R-module. Let $m_1, m_2, \ldots, m_n$ be finitely many vectors in M.

(a) A **linear combination** of $m_1, m_2, \ldots, m_n$ means a vector of the form

$$r_1 m_1 + r_2 m_2 + \cdots + r_n m_n \qquad \text{with } r_1, r_2, \ldots, r_n \in R.$$

(b) The set of all linear combinations of $m_1, m_2, \ldots, m_n$ is called the **span** of $(m_1, m_2, \ldots, m_n)$, and is denoted by $\operatorname{span}(m_1, m_2, \ldots, m_n)$. (Note that [DF] calls it $R\{m_1, m_2, \ldots, m_n\}$.)

(c) If the span of $(m_1, m_2, \ldots, m_n)$ is M, then we say that the vectors $m_1, m_2, \ldots, m_n$ **span** M (or **generate** M).

(d) We say that the vectors $m_1, m_2, \ldots, m_n$ are **linearly independent** if the following holds: If $r_1, r_2, \ldots, r_n \in R$ satisfy

$$r_1 m_1 + r_2 m_2 + \cdots + r_n m_n = 0,$$

then $r_1 = r_2 = \cdots = r_n = 0$. (In other words, the vectors $m_1, m_2, \ldots, m_n$ are said to be linearly independent if the only way to write 0 as a linear combination of them is $0 = 0m_1 + 0m_2 + \cdots + 0m_n$.)

(e) We say that the n-tuple $(m_1, m_2, \ldots, m_n)$ is a **basis** of the R-module M if $m_1, m_2, \ldots, m_n$ are linearly independent and span M.

(f) All of this terminology depends on R. Thus, if R is not clear from the context, we will clarify it by saying "R-linear combination" (or "linear combination over R") instead of just "linear combination", and likewise saying "R-span" or "R-linearly independent" or "R-basis".

Fine print: The property of n vectors $m_1, m_2, \ldots, m_n$ to span M is a joint property (i.e., it is a property of the **list** $(m_1, m_2, \ldots, m_n)$, not of each single vector). The same applies to linear independence. Sometimes, we do say that a single vector m spans M (for example, the vector $1 \in \mathbb{Z}$ spans the $\mathbb{Z}$-module $\mathbb{Z}$); this means that the one-element list (m) spans M.

Definition 1.2.1 was tailored to finite lists of vectors, but we can extend it to arbitrary (possibly infinite) families of vectors:

Definition 1.2.2. Let M be a left R-module. Let $(m_i)_{i \in I}$ be a family of vectors in M (with I being any set).

(a) A **linear combination** of $(m_i)_{i \in I}$ means a vector of the form

$$\sum_{i \in I} r_i m_i$$

for some family $(r_i)_{i \in I}$ of scalars (i.e., for some choice of $r_i \in R$ for each $i \in I$) with the property that

$$\text{all but finitely many } i \in I \text{ satisfy } r_i = 0. \tag{2}$$

Here, the sum $\sum\limits_{i \in I} r_i m_i$ is an infinite sum, but all but finitely many of its addends are zero (thanks to the condition (2)). Such a sum is simply defined to be the sum of the nonzero addends. For example, $3 + 2 + 0 + 0 + 0 + \cdots = 3 + 2 = 5$.

(b) The set of all linear combinations of $(m_i)_{i \in I}$ is called the **span** of $(m_i)_{i \in I}$, and is denoted by $\operatorname{span}(m_i)_{i \in I}$. (Note that [DF] calls it $R\{m_i \mid i \in I\}$.)

(c) If the span of $(m_i)_{i \in I}$ is M, then we say that the family $(m_i)_{i \in I}$ **spans** M (or **generates** M).

(d) We say that the family $(m_i)_{i \in I}$ is **linearly independent** if the following holds: If $r_i \in R$ satisfy

$$\text{all but finitely many } i \in I \text{ satisfy } r_i = 0 \tag{3}$$

and

$$\sum_{i \in I} r_i m_i = 0,$$

then $r_i = 0$ for all $i \in I$.

(e) We say that the family $(m_i)_{i \in I}$ is a **basis** of the R-module M if $(m_i)_{i \in I}$ is linearly independent and spans M.

(f) All of this terminology depends on R. Thus, if R is not clear from the context, we will clarify it by saying "R-linear combination" (or "linear combination over R") instead of just "linear combination", etc..

The infinite sums in this definition are a bit of a distraction, but a necessary one. Fortunately, when studying these notions, it is often sufficient to work with finite families (i.e., finite sets I), since they are in some sense representative of the general case. To wit:

Proposition 1.2.3. Let M be a left R-module. Let $(m_i)_{i \in I}$ be a family of vectors in M (with I being any set).

(a) Any linear combination of $(m_i)_{i \in I}$ is already a linear combination of some finite subfamily of $(m_i)_{i \in I}$. (That is: If m is a linear combination of $(m_i)_{i \in I}$, then there exists some finite subset J of I such that m is a linear combination of $(m_i)_{i \in J}$.)

(b) The family $(m_i)_{i \in I}$ is linearly independent if and only if all its finite subfamilies (i.e., all families of the form $(m_i)_{i \in J}$ with J being a finite subset of I) are linearly independent.

Proof. **(a)** Let m be a linear combination of $(m_i)_{i \in I}$. Thus, m has the form

$$m = \sum_{i \in I} r_i m_i$$

for some family $(r_i)_{i \in I}$ of scalars (i.e., for some choice of $r_i \in R$ for each $i \in I$) with the property that

all but finitely many $i \in I$ satisfy $r_i = 0$.

The latter property can be rewritten as follows: There exists a **finite** subset J of I such that all $i \in I \setminus J$ satisfy $r_i = 0$. Consider this J. Then, in the sum $\sum\limits_{i \in I} r_i m_i$, all the addends with $i \notin J$ are 0 (since these addends satisfy $i \notin J$, thus $i \in I \setminus J$, hence $r_i = 0$ and therefore $r_i m_i = 0 m_i = 0$). Hence, we can throw these addends away and are left with the finite sum $\sum\limits_{i \in J} r_i m_i$. Therefore, $\sum\limits_{i \in I} r_i m_i = \sum\limits_{i \in J} r_i m_i$, so that $m = \sum\limits_{i \in I} r_i m_i = \sum\limits_{i \in J} r_i m_i$. This shows that m is a linear combination of the finite subfamily $(m_i)_{i \in J}$ of our original family $(m_i)_{i \in I}$. This proves Proposition 1.2.3 **(a)**.

(b) This is similar to part **(a)**. The details are left to the reader. (Again, the key is that the condition (3) allows us to restrict ourselves to a finite subset of I.) □

Next, we show that the span of a family of vectors is always a submodule:

Proposition 1.2.4. Let M be a left R-module. Let $(m_i)_{i \in I}$ be a family of vectors in M. Then, the span of this family is an R-submodule of M.

Proof. You have to show the following three statements:

1. The sum of two linear combinations of $(m_i)_{i \in I}$ is a linear combination of $(m_i)_{i \in I}$.

2. Scaling a linear combination of $(m_i)_{i\in I}$ by an $r \in R$ gives a linear combination of $(m_i)_{i\in I}$.

3. The zero vector is a linear combination of $(m_i)_{i\in I}$.

All three of these statements are easy. For example, let me show the first statement: Let v and w be two linear combinations of $(m_i)_{i\in I}$. Thus, we can write v and w as

$$v = \sum_{i\in I} a_i m_i \qquad \text{and} \qquad w = \sum_{i\in I} b_i m_i \tag{4}$$

for some two families $(a_i)_{i\in I}$ and $(b_i)_{i\in I}$ of scalars (i.e., for some choices of $a_i \in R$ and $b_i \in R$ for each $i \in I$) with the property that

$$\text{all but finitely many } i \in I \text{ satisfy } a_i = 0 \tag{5}$$

and that

$$\text{all but finitely many } i \in I \text{ satisfy } b_i = 0. \tag{6}$$

Now, adding the two equalities in (4) together, we obtain

$$\begin{aligned} v + w &= \sum_{i\in I} a_i m_i + \sum_{i\in I} b_i m_i = \sum_{i\in I} (a_i m_i + b_i m_i) \\ &= \sum_{i\in I} (a_i + b_i) m_i. \end{aligned} \tag{7}$$

Moreover, combining (5) with (6), we see that all but finitely many $i \in I$ satisfy $a_i + b_i = 0$ (since the union of two finite sets is still a finite set). Hence, (7) shows that $v + w$ is a linear combination of $(m_i)_{i\in I}$. This proves Statement 1 above. The proofs of Statements 2 and 3 are even easier. □

Definition 1.2.5. **(a)** A left R-module is said to be **free** if it has a basis.

(b) Let $n \in \mathbb{N}$. A left R-module is said to be **free of rank** n if it has a basis of size n (i.e., a basis consisting of n vectors).

Let us see some examples of modules that are free and modules that aren't.

You might want to look at $\mathbb{Q}$-modules at first; but they make for boring examples, because of the following fact:

Theorem 1.2.6. If F is a field, then every F-module (= F-vector space) is free.

Proof. This is just the famous fact from linear algebra that every vector space has a basis. In the most important case (which is when the vector space admits a finite spanning set – i.e., there is a finite list $(m_1, m_2, \ldots, m_n)$ of vectors that spans it[1]), this has fairly neat elementary proofs (see, e.g., Theorem 2.1 in Keith Conrad's `https://kconrad.math.uconn.edu/blurbs/linmultialg/dimension.pdf`). In the general case, the proof is tricky and requires the Axiom of Choice (see Theorem 4.1 in Keith Conrad's `https://kconrad.math.uconn.edu/blurbs/zorn1.pdf`). □

[1] Such vector spaces are called **finite-dimensional**.

For example, Theorem 1.2.6 shows that the $\mathbb{Q}$-vector space $\mathbb{R}$ is free, i.e., has a basis. Such bases are called **Hamel bases** and theoretically exist (if you believe in the Axiom of Choice). Practically, there is no way to construct one.

To find more interesting examples, we need to consider rings that are not fields. First of all, let us discuss a family of examples that exists for an arbitrary ring R:

- Consider the left R-module

  $$R^2 = \{(a,b) \mid a \in R \text{ and } b \in R\}.$$

  This R-module R^2 is free of rank 2, since the list $((1,0), \ (0,1))$ is a basis of it. Indeed:

  - The vectors $(1,0), \ (0,1)$ span R^2 (because any vector (a,b) can be written as $a(1,0) + b(0,1)$, and thus is a linear combination of $(1,0), \ (0,1)$).
  - The vectors $(1,0), \ (0,1)$ are linearly independent, since $a(1,0) + b(0,1) = (a,b)$ can only be 0 if $a = b = 0$.

- Likewise, the left R-module R^3 has basis $((1,0,0), \ (0,1,0), \ (0,0,1))$.

- More generally: If $n \in \mathbb{N}$, then the left R-module R^n has basis

  $$\begin{aligned}((1,0,0,\ldots,0)&,\\ (0,1,0,\ldots,0)&,\\ (0,0,1,\ldots,0)&,\\ \ldots&,\\ (0,0,0,\ldots,1)&).\end{aligned}$$

  This basis is called the **standard basis** of R^n, and its n vectors are called $e_1, e_2, \ldots, e_n$ (in this order). To make this more rigorous: For each $i \in \{1, 2, \ldots, n\}$, we define e_i to be the vector in R^n whose i-th entry is 1 and whose all remaining entries are 0 (it is an n-tuple, like any vector in R^n). Then, the list $(e_1, e_2, \ldots, e_n)$ is a basis of the left R-module R^n. Thus, the R-module R^n is free of rank n.

- As a particular case, the left R-module R^1 is free of rank 1. Note that $R^1 \cong R$, because the map $R \to R^1$, $r \mapsto (r)$ (which merely wraps each scalar into a list to turn it into a vector) is an R-module isomorphism. Hence, the left R-module R is free of rank 1. Of course, you can see this directly as well: The one-element list (1) is a basis of it.

  Likewise, the left R-module R^0 is free of rank 0. Note that R^0 is a trivial R-module (it consists of just the zero vector); the empty list is a basis for it (since the only vector in R^0 is the zero vector and thus is a linear combination of nothing). Some authors (e.g., Keith Conrad in the abovementioned references) avoid trivial R-modules[2], but there is no natural reason to do so except for the slight weirdness of dealing with empty lists and empty sums.

- More generally: If I is a set, then $R^{(I)} = \bigoplus_{i \in I} R$ is a free R-module. It has a standard basis $(e_i)_{i \in I}$, where each e_j is a family that has a 1 in its j-th position and 0s in all other positions. (That is, $e_j = \left(\delta_{i,j}\right)_{i \in I}$, where $\delta_{i,j} = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } i \neq j \end{cases}$.)

  In general, the R-module $R^I = \prod_{i \in I} R$ is not free. (For example, the $\mathbb{Z}$-module $\mathbb{Z}^{\mathbb{N}}$ is not free. This is actually not easy to prove! A proof is sketched in [DF, §10.3, Exercise 24]. It is easy to see that the standard basis $(e_i)_{i \in \mathbb{N}}$ of $\mathbb{Z}^{(\mathbb{N})}$ is not a basis of $\mathbb{Z}^{\mathbb{N}}$, since (e.g.) the vector $(1, 1, 1, 1, \ldots)$ is not a linear combination of this family[3]. But it is much harder to show that there is no basis at all.)

Let us now look at $\mathbb{Z}$-modules. Recall that $\mathbb{Z}$-modules are the same as abelian groups, so free $\mathbb{Z}$-modules are also known as **free abelian groups** (this is not the same as free groups).

- Consider the $\mathbb{Z}$-submodule
$$U := \left\{ (a, b, c) \in \mathbb{Z}^3 \ \mid \ a + b + c = 0 \right\} \text{ of } \mathbb{Z}^3.$$

  Is U free? Can we find a basis for U ?

  So we are trying to find a basis for a submodule of $\mathbb{Z}^3$ that is determined by a set of linear equations (in our case, only one linear equation – namely, $a + b + c = 0$). If we were using a field (e.g., $\mathbb{Q}$ or $\mathbb{R}$) instead of $\mathbb{Z}$, then this would be a familiar problem (equivalent to solving a system of homogeneous linear equations), which can be solved by Gaussian elimination. If we try to perform Gaussian elimination over $\mathbb{Z}$, we might run into trouble: Denominators may appear; as a result, we might not actually get vectors with integer entries. However, for the U above, this does not happen, and we get the basis
$$((-1, 0, 1), \ (0, -1, 1)).$$

[2] A **trivial R-module** means an R-module that consists only of the zero vector.

[3] Of course, you could write
$$(1, 1, 1, 1, \ldots) = 1e_0 + 1e_1 + 1e_2 + 1e_3 + 1e_4 + \cdots;$$
however, the sum on the right is properly infinite (with infinitely many nonzero coefficients) and thus does not count as a linear combination (as it fails the condition (2) from Definition 1.2.2).

So U is indeed free.

What if we have a more complicated submodule and we do run into denominators? Thus, we do not get a basis using Gaussian elimination. Does this mean that no basis exists, or does it mean we have to try something else? We will soon see.

- The $\mathbb{Z}$-module $\mathbb{Z}/2$ is not free (i.e., does not have a basis). Indeed, if it had a basis, then this basis would contain at least one vector (since $\mathbb{Z}/2$ is not trivial), but this vector would not be linearly independent, since scaling it by 2 would give 0.

- The $\mathbb{Z}$-module $\mathbb{Q}$ is not free (i.e., does not have a basis).

  Proof. Assume the contrary. Thus, there exists a $\mathbb{Z}$-basis $(m_i)_{i\in I}$ of $\mathbb{Q}$. The set I must be nonempty (since $\mathbb{Q}$ is not trivial); thus, we are in one of the following two cases:

  - *Case 1:* We have $|I| = 1$. In this case, I is a 1-element set, so we can rewrite our basis $(m_i)_{i\in I}$ as a list (m) that consists of a single rational number m. This single rational number m must span the entire $\mathbb{Z}$-module $\mathbb{Q}$. In other words, every element of $\mathbb{Q}$ must be a $\mathbb{Z}$-multiple of m. But this is absurd (indeed, if $m = 0$, then 1 is not a $\mathbb{Z}$-multiple of m; but otherwise, $\dfrac{1}{2}m$ is not a $\mathbb{Z}$-multiple of m).
  - *Case 2:* We have $|I| > 1$. In this case, there are at least two vectors m_u and m_v in this basis $(m_i)_{i\in I}$. However, two rational numbers are never $\mathbb{Z}$-linearly independent[4]. Thus, a fortiori, the whole family $(m_i)_{i\in I}$ cannot be $\mathbb{Z}$-linearly independent (since a subfamily of a linearly independent family of vectors must always be linearly independent). This contradicts the assumption that this family is a basis.

  Thus, in each case, we have found a contradiction, and our proof is complete.

- Now, consider the $\mathbb{Z}$-submodule

$$V := \left\{(a,b) \in \mathbb{Z}^2 \mid a \equiv b \bmod 2\right\} \text{ of } \mathbb{Z}^2.$$

[4] Indeed, let p and q be two rational numbers. We claim that there exist integers $a, b \in \mathbb{Z}$ that are not both 0 but still satisfy $ap + bq = 0$. (This will clearly prove that p and q are not $\mathbb{Z}$-linearly independent.)

Indeed, if $p = 0$, then we set $a = 1$ and $b = 0$ and are done. Something similar works if $q = 0$. So we WLOG assume that $p \neq 0$ and $q \neq 0$. Write p and q as $p = \dfrac{n}{d}$ and $q = \dfrac{m}{e}$ for some nonzero integers n, d, m, e (we can do this, since p and q are nonzero rational numbers). Then, $dmp + (-en)q = 0$ (check this!), so we have found our a and b (namely, $a = dm$ and $b = -en$).

This $\mathbb{Z}$-submodule V contains the vectors $(0,2)$ and $(1,1)$ and $(1,-1)$ and $(4,-2)$ and many others. Is V free? Can we find a basis for V ?

Let's try the pair $((2,0), \ (0,2))$. Is this pair a basis for V ? Its span is

$$\begin{aligned}\operatorname{span}((2,0), \ (0,2)) &= \{c\,(2,0) + d\,(0,2) \ \mid \ c,d \in \mathbb{Z}\} \\ &= \{(2c,2d) \ \mid \ a,b \in \mathbb{Z}\} \\ &= \left\{(a,b) \in \mathbb{Z}^2 \ \mid \ a \equiv b \equiv 0 \bmod 2\right\}.\end{aligned}$$

This is a $\mathbb{Z}$-submodule of V, but not the entire V, since (for example) $(1,1)$ belongs to V but not to $\operatorname{span}((2,0), \ (0,2))$. So we have "undershot" our V (by finding a linearly independent family that does not span V).

Let's try the triple $((2,0), \ (0,2), \ (1,1))$. This triple does span V (check this!), but is not linearly independent, since

$$1 \cdot (2,0) + 1 \cdot (0,2) + (-2) \cdot (1,1) = 0.$$

So we have "overshot" V now (by finding a family that spans V but is not linearly independent).

Let us try to correct this by throwing away $(0,2)$. So we are left with the pair $((2,0), \ (1,1))$. And this pair is indeed a basis of V, as can easily be checked. Indeed, it is linearly independent (you can check this using linear algebra, since it clearly suffices to prove its $\mathbb{Q}$-linear independence), and furthermore spans V because each $(a,b) \in V$ can be written as a linear combination of $(2,0), \ (1,1)$ as follows:

$$(a,b) = \underbrace{\frac{a-b}{2}}_{\substack{\in\mathbb{Z} \\ \text{(since } a\equiv b \bmod 2)}} \cdot (2,0) + b \cdot (1,1).$$

Another basis for V is the pair $((1,1), \ (1,-1))$. Indeed, this pair is linearly independent (check this!), and it spans V, because each $(a,b) \in V$ can be written as

$$(a,b) = \underbrace{\frac{a+b}{2}}_{\in\mathbb{Z}} \cdot (1,1) + \underbrace{\frac{a-b}{2}}_{\in\mathbb{Z}} \cdot (1,-1) \in \operatorname{span}((1,1), \ (1,-1)).$$

Let us now return to the general case to state a few theorems:

Theorem 1.2.7. Let M be a left R-module. Let $n \in \mathbb{N}$. The left R-module M is free of rank n if and only if $M \cong R^n$ (as left R-modules).

More concretely:

Theorem 1.2.8. Let M be a left R-module. Let $m_1, m_2, \ldots, m_n$ be n vectors in M. Consider the map

$$\begin{aligned} f : R^n &\to M, \\ (r_1, r_2, \ldots, r_n) &\mapsto r_1 m_1 + r_2 m_2 + \cdots + r_n m_n. \end{aligned}$$

Then:

(a) This map f is always a left R-module morphism.

(b) The map f is injective if and only if $m_1, m_2, \ldots, m_n$ are linearly independent.

(c) The map f is surjective if and only if $m_1, m_2, \ldots, m_n$ span M.

(d) The map f is an isomorphism[5] if and only if $(m_1, m_2, \ldots, m_n)$ is a basis of M.

Note that the map f in Theorem 1.2.8 takes an n-tuple $(r_1, r_2, \ldots, r_n)$ of scalars, and uses these scalars as coefficients to form a linear combination of $m_1, m_2, \ldots, m_n$. Thus, the values of f are precisely the linear combinations of $m_1, m_2, \ldots, m_n$.

Proof of Theorem 1.2.8. This is commonly done in linear algebra texts (albeit usually under the assumption that R is a field, but the proof is the same); thus I will be brief.

(a) We must prove that f respects addition, respects scaling and respects the zero. I will only show that it respects addition, since the other two statements are analogous.

So we must prove that $f(a+b) = f(a) + f(b)$ for all $a, b \in R^n$. Indeed, let $a, b \in R^n$. Write a and b as

$$a = (a_1, a_2, \ldots, a_n) \qquad \text{and} \qquad b = (b_1, b_2, \ldots, b_n).$$

Then, the definition of R^n (as the direct product $\underbrace{R \times R \times \cdots \times R}_{n \text{ times}}$) yields $a + b = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$. Hence, the definition of f yields

$$\begin{aligned} f(a+b) &= (a_1 + b_1) m_1 + (a_2 + b_2) m_2 + \cdots + (a_n + b_n) m_n \\ &= (a_1 m_1 + b_1 m_1) + (a_2 m_2 + b_2 m_2) + \cdots + (a_n m_n + b_n m_n) \\ &\qquad \text{(by right distributivity)} \\ &= \underbrace{(a_1 m_1 + a_2 m_2 + \cdots + a_n m_n)}_{\substack{= f(a) \\ \text{(by the definition of } f \text{, since } a = (a_1, a_2, \ldots, a_n))}} + \underbrace{(b_1 m_1 + b_2 m_2 + \cdots + b_n m_n)}_{\substack{= f(b) \\ \text{(by the definition of } f \text{, since } b = (b_1, b_2, \ldots, b_n))}} \\ &= f(a) + f(b), \end{aligned}$$

which is what we wanted to show.

[5] Of course, "isomorphism" means "left R-module isomorphism" here.

(b) The map f is an R-module morphism (by part **(a)**). Thus, it is injective if and only if $\operatorname{Ker} f = \{0_{R^n}\}$ (by Lemma 1.1.8). Hence, we have the following chain of logical equivalences:

$$\begin{aligned}
& (f \text{ is injective}) \\
\Longleftrightarrow\ & (\operatorname{Ker} f = \{0_{R^n}\}) \\
\Longleftrightarrow\ & (\operatorname{Ker} f \subseteq \{0_{R^n}\}) \qquad (\text{since } \{0_{R^n}\} \text{ is clearly a subset of } \operatorname{Ker} f) \\
\Longleftrightarrow\ & (\{a \in R^n \mid f(a) = 0\} \subseteq \{0_{R^n}\}) \\
& \qquad\qquad (\text{since } \operatorname{Ker} f = \{a \in R^n \mid f(a) = 0\} \text{ by the definition of } \operatorname{Ker} f) \\
\Longleftrightarrow\ & (\text{the only } a \in R^n \text{ satisfying } f(a) = 0 \text{ is } 0_{R^n}) \\
\Longleftrightarrow\ & \left(\begin{array}{c} \text{the only } (a_1, a_2, \ldots, a_n) \in R^n \text{ satisfying } f(a_1, a_2, \ldots, a_n) = 0 \\ \text{is } (0, 0, \ldots, 0) \end{array}\right) \\
& \qquad\qquad \left(\begin{array}{c} \text{since any } a \in R^n \text{ can be written in the form } (a_1, a_2, \ldots, a_n), \\ \text{and since } 0_{R^n} = (0, 0, \ldots, 0) \end{array}\right) \\
\Longleftrightarrow\ & \left(\begin{array}{c} \text{the only } (a_1, a_2, \ldots, a_n) \in R^n \text{ satisfying } a_1 m_1 + a_2 m_2 + \cdots + a_n m_n = 0 \\ \text{is } (0, 0, \ldots, 0) \end{array}\right) \\
& \qquad\qquad \left(\begin{array}{c} \text{since } f(a_1, a_2, \ldots, a_n) = a_1 m_1 + a_2 m_2 + \cdots + a_n m_n \\ \text{for any } (a_1, a_2, \ldots, a_n) \in R^n \end{array}\right) \\
\Longleftrightarrow\ & \left(\begin{array}{c} \text{if } a_1, a_2, \ldots, a_n \in R \text{ satisfy } a_1 m_1 + a_2 m_2 + \cdots + a_n m_n = 0, \\ \text{then } a_1 = a_2 = \cdots = a_n = 0 \end{array}\right) \\
\Longleftrightarrow\ & (m_1, m_2, \ldots, m_n \text{ are linearly independent})
\end{aligned}$$

(by the definition of linear independence). This proves part **(b)** of the theorem.

(c) We have the following chain of logical equivalences:

$$\begin{aligned}
& (f \text{ is surjective}) \\
\Longleftrightarrow\ & (\text{each } m \in M \text{ can be written as } f(a) \text{ for some } a \in R^n) \\
\Longleftrightarrow\ & \left(\begin{array}{c} \text{each } m \in M \text{ can be written as } f(a_1, a_2, \ldots, a_n) \\ \text{for some } (a_1, a_2, \ldots, a_n) \in R^n \end{array}\right) \\
& \qquad\qquad (\text{since any } a \in R^n \text{ can be written in the form } (a_1, a_2, \ldots, a_n)) \\
\Longleftrightarrow\ & \left(\begin{array}{c} \text{each } m \in M \text{ can be written as } a_1 m_1 + a_2 m_2 + \cdots + a_n m_n \\ \text{for some } (a_1, a_2, \ldots, a_n) \in R^n \end{array}\right) \\
& \qquad\qquad \left(\begin{array}{c} \text{since } f(a_1, a_2, \ldots, a_n) = a_1 m_1 + a_2 m_2 + \cdots + a_n m_n \\ \text{for any } (a_1, a_2, \ldots, a_n) \in R^n \end{array}\right) \\
\Longleftrightarrow\ & (\text{each } m \in M \text{ is a linear combination of } m_1, m_2, \ldots, m_n) \\
& \qquad\qquad (\text{by the definition of a linear combination}) \\
\Longleftrightarrow\ & (m_1, m_2, \ldots, m_n \text{ span } M).
\end{aligned}$$

This proves part **(c)** of the theorem.

(d) We have the following chain of logical equivalences:

$$\begin{aligned}
&(f \text{ is an } R\text{-module isomorphism})\\
\Longleftrightarrow\ &(f \text{ is invertible})\\
&\qquad\qquad\left(\begin{array}{c}\text{since we know from Proposition 1.1.1 that any}\\ \text{invertible } R\text{-module morphism is an isomorphism}\end{array}\right)\\
\Longleftrightarrow\ &(f \text{ is bijective})\\
\Longleftrightarrow\ &\qquad \underbrace{(f \text{ is injective})}_{\substack{\Longleftrightarrow\ (m_1,m_2,\ldots,m_n \text{ are linearly independent})\\ (\text{by part } \textbf{(b)})}} \quad\wedge\quad \underbrace{(f \text{ is surjective})}_{\substack{\Longleftrightarrow\ (m_1,m_2,\ldots,m_n \text{ span } M)\\ (\text{by part } \textbf{(c)})}}\\
\Longleftrightarrow\ &(m_1,m_2,\ldots,m_n \text{ are linearly independent}) \wedge (m_1,m_2,\ldots,m_n \text{ span } M)\\
\Longleftrightarrow\ &((m_1,m_2,\ldots,m_n) \text{ is a basis of } M)
\end{aligned}$$

(by the definition of a basis). This proves part **(d)** of the theorem. □

Proof of Theorem 1.2.7. $\Longrightarrow$: Assume that M is free of rank n. That is, M has a basis $(m_1, m_2, \ldots, m_n)$ of size n. Consider this basis. Consider the map $f : R^n \to M$ defined in Theorem 1.2.8. Thus, Theorem 1.2.8 **(d)** yields that f is an isomorphism. Hence, $R^n \cong M$ as left R-modules. In other words, $M \cong R^n$ as left R-modules. This proves the "$\Longrightarrow$" direction of Theorem 1.2.7.

$\Longleftarrow$: Assume that $M \cong R^n$ as left R-modules. But the left R-module R^n is free of rank n (as we have seen above). Hence, I claim that the left R-module M is also free of rank n, since $M \cong R^n$. Indeed, this follows from the "meta-theorem" that says that module isomorphisms preserve all "intrinsic" properties of modules (in this case, this property is "being free of rank n").

Here is a more pedestrian way to get to the same conclusion: We have $M \cong R^n$, thus $R^n \cong M$. In other words, there exists a left R-module isomorphism $g : R^n \to M$. Consider this g. Now, consider the standard basis $(e_1, e_2, \ldots, e_n)$ of the left R-module R^n. Applying g to each vector in this basis, we obtain a list $(g(e_1), g(e_2), \ldots, g(e_n))$ of vectors in M. It is straightforward to see that this new list is a basis of M (indeed, when we apply g to a linear combination $a_1e_1 + a_2e_2 + \cdots + a_ne_n$ of the standard basis $(e_1, e_2, \ldots, e_n)$ in R^n, then we obtain

$$g(a_1e_1 + a_2e_2 + \cdots + a_ne_n) = a_1g(e_1) + a_2g(e_2) + \cdots + a_ng(e_n) \qquad (\text{since } g \text{ is } R\text{-linear}),$$

which is the corresponding linear combination of $(g(e_1), g(e_2), \ldots, g(e_n))$; thus, linear independence of $(e_1, e_2, \ldots, e_n)$ translates into linear independence of $(g(e_1), g(e_2), \ldots, g(e_n))$ (since g sends only 0 to 0), and the same holds for spanning (since g is bijective)). Hence, M has a basis of size n. In other words, M is free of rank n.

Either way, the "$\Longleftarrow$" direction of Theorem 1.2.7 is now proved. □

Theorem 1.2.8 can be generalized to bases of arbitrary size:

Theorem 1.2.9. Let M be a left R-module. Let $(m_i)_{i\in I}$ be any family of vectors in M. Consider the map

$$f : R^{(I)} \to M,$$
$$(r_i)_{i\in I} \mapsto \sum_{i\in I} r_i m_i.$$

(This is well-defined, since any $(r_i)_{i\in I} \in R^{(I)}$ automatically satisfies the condition (2) because of the definition of $R^{(I)}$.)

Then:

(a) This map f is always a left R-module morphism.

(b) The map f is injective if and only if the family $(m_i)_{i\in I}$ is linearly independent.

(c) The map f is surjective if and only if the family $(m_i)_{i\in I}$ spans M.

(d) The map f is an isomorphism if and only if the family $(m_i)_{i\in I}$ is a basis of M.

Note that the map f here has domain $R^{(I)}$, not R^I, since the infinite sum $\sum_{i\in I} r_i m_i$ is well-defined for all $(r_i)_{i\in I} \in R^{(I)}$ but not (in general) for all $(r_i)_{i\in I} \in R^I$.

Proof of Theorem 1.2.9. Analogous to Theorem 1.2.8, with the usual allowance for infinite sums. □

Remark 1.2.10. As you will have noticed by now, "free module of rank n" is a generalization of "vector space of dimension n" to arbitrary rings.

We have been careful to speak of "free modules of rank n", but never of "the rank of a free module". This is due to the somewhat perverse-sounding fact that there can be modules that are free of several ranks simultaneously (i.e., modules that have bases of different sizes). One way to get such modules is by taking R to be a trivial ring (in which case, any R-module is trivial and is free of every rank simultaneously – seriously). If this was the only example, one could discount the issue as a formality, but there are less trivial (pardon) examples as well: [DF, §10.3, exercise 27] constructs a ring R over which $R^n \cong R$ as left R-modules for each $n \in \{1, 2, 3, \ldots\}$ (so R itself is a free R-module of rank n for each $n \in \{1, 2, 3, \ldots\}$).

If R is a nontrivial commutative ring, then things are nice: The R-modules $R^0, R^1, R^2, \ldots$ are mutually non-isomorphic, so a free R-module can never have two different ranks at the same time. This is not obvious (see [DF, §10.3, exercise 2]). We can actually say more: If R is a nontrivial commutative ring, then an R-module morphism $R^m \to R^n$ cannot be injective unless $m \leq n$ (see, e.g., `https://math.stackexchange.com/questions/106786`), and cannot be surjective unless $m \geq n$ (see, e.g., `https://math.stackexchange.com/questions/20178`). These facts are in line with the intuition you should have from linear algebra (injective maps cannot quash dimensions; surjective maps cannot create dimensions) and also with the Pigeonhole Principles from combinatorics (a map between two finite sets M and N cannot be injective unless $|M| \leq |N|$, and cannot be surjective unless $|M| \geq |N|$). But actually proving them takes real work!