Homological Algebra

Math 915 Fall 2023

Eloísa Grifo

August 1, 2026

Warning!

Proceed with caution. These notes are under construction and are 100% guaranteed to contain typos. If you find any typos or errors, I will be most grateful to you for letting me know. If you are looking for a place where to learn homological algebra or category theory, I strongly recommend the following excellent resources:

- Rotman's *An introduction to homological algebra*, second edition. [Rot09]
- Weibel's *Homological Algebra* [Wei94].
- Mac Lane's *Categories for the working mathematician* [ML98].
- Emily Riehl's *Category Theory in context*.

Acknowledgements

These notes are partially based on notes I wrote in Spring 2021 for a homological algebra class at the University of California, Riverside, and draw heavily from the references above. These notes also owe a big debt to all the students in that class for their comments and questions that lead to multiple improvements, especially to those who found typos: Brandon Massaro, Rahul Rajkumar, Adam Richardson, Khoa Ta, Ryan Watson, and Noble Williamson. Thank you also to those who found typos in this new version: Stephen Stern, Danny Anderson, Taylor Murray, Kesavan M, Robert Ireland, Kolton O'Neal, Sam Macdonald, Alex Heitzman, Sara Mueller, Sabrina Klement, Deron Lessure, Julie Geraci, Siddhartha Pramanik, Debojyoti Pal, Danny Anderson, Nicole Xie, and Em Stephen.

A special thanks to Luis Núñez Betancourt, who used these notes in his Spring 2026 course, for all his helpful comments and suggestions.

Contents

Chapter 0

Where are we going?

Homological algebra first appeared in the study of topological spaces. Roughly speaking, homology is a way of associating a sequence of abelian groups (or modules, or other more sophisticated algebraic objects) to another object, for example a topological space. The homology of a topological space encodes topological information about the space in algebraic language — this is what algebraic topology is all about.

More formally, we will study *complexes* and their homology from a more abstract perspective. While algebraic topologists are often concerned with complexes of abelian groups, we will work a bit more generally with complexes of R-modules. The basic assumptions and notation about rings and modules we will use in this class can be found in Appendix A. As an appetizer, we begin with some basic homological algebra definitions.

Definition 0.1. A **chain complex** of R-modules $(C_\bullet, \partial_\bullet)$, also referred to simply as a **complex**, is a sequence of R-modules C_i and R-module homomorphisms

$$\cdots \longrightarrow C_{n+1} \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_n} C_{n-1} \longrightarrow \cdots$$

such that $\partial_n \partial_{n+1} = 0$ for all n. We refer to C_n as the module in **homological degree** n. The maps ∂_n are the **differentials** of our complex. We may sometimes omit the differentials ∂_n and simply refer to the complex $C_\bullet$ or even C; we may also sometimes refer to $\partial_\bullet$ as *the* differential of $C_\bullet$.

In some contexts, it is important to make a distinction between chain complexes and co-chain complexes, where the arrows go the opposite way: a co-chain complex would look like

$$\cdots \longrightarrow C_{n-1} \xrightarrow{\partial_n} C_n \xrightarrow{\partial_{n+1}} C_{n+1} \longrightarrow \cdots .$$

We will not need to make such a distinction, so we will call both of these complexes and most often follow the convention in the definition above. We will say a complex C is **bounded above** if $C_n = 0$ for all $n \gg 0$, and **bounded below** if $C_n = 0$ for all $n \ll 0$. A **bounded complex** is one that is both bounded above and below. If a complex is bounded, we may sometimes simply write it as a finite complex, say

$$C_n \xrightarrow{\partial_n} C_{n-1} \longrightarrow \cdots \longrightarrow C_m.$$

Remark 0.2. The condition that $\partial_n \partial_{n+1} = 0$ for all n implies that $\operatorname{im} \partial_{n+1} \subseteq \ker \partial_n$.

Definition 0.3. The complex $(C_\bullet, \partial_\bullet)$ is **exact** at n if $\operatorname{im} \partial_{n+1} = \ker \partial_n$. An **exact sequence** is a complex that is exact everywhere. More precisely, an **exact sequence** of R-modules is a sequence

$$\cdots \xrightarrow{f_{n-1}} M_n \xrightarrow{f_n} M_{n+1} \xrightarrow{f_{n+1}} \cdots$$

of R-modules and R-module homomorphisms such that $\operatorname{im} f_n = \ker f_{n+1}$ for all n. An exact sequence of the form

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

is a **short exact sequence**, sometimes written **ses**.

Remark 0.4. The sequence

$$0 \longrightarrow M \xrightarrow{f} N$$

is exact if and only if f is injective. Similarly,

$$M \xrightarrow{f} N \longrightarrow 0$$

is exact if and only if f is surjective. So

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

is a short exact sequence if and only if

- f is injective
- g is surjective
- $\operatorname{im} f = \ker g$.

When this is indeed a short exact sequence, we can identify A with its image $f(A)$, and $A = \ker g$. Moreover, since g is surjective, by the First Isomorphism Theorem we conclude that $C \cong B/f(A)$, so we might abuse notation and identify C with B/A.

Notation 0.5. We write $A \twoheadrightarrow B$ to denote a surjective map, and $A \hookrightarrow B$ to denote an injective map.

Definition 0.6. The **cokernel** of a map of R-modules $A \xrightarrow{f} B$ is the module

$$\operatorname{coker} f := B/\operatorname{im}(f).$$

Remark 0.7. We can rephrase Theorem 0.4 in a fancier language: if

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

is a short exact sequence, then $A = \ker g$ and $C = \operatorname{coker} f$.

Example 0.8. Let π be the canonical projection $\mathbb{Z} \longrightarrow \mathbb{Z}/2\mathbb{Z}$. The following is a short exact sequence:

$$0 \longrightarrow \mathbb{Z} \xrightarrow{2} \mathbb{Z} \xrightarrow{\pi} \mathbb{Z}/2\mathbb{Z} \longrightarrow 0\,.$$

We will most often be interested in **complexes of R-modules**, where the abelian groups that show up are all modules over the same ring R.

Example 0.9. Let $R = k[x]$ be a polynomial ring over the field k. The following is a short exact sequence:

$$0 \longrightarrow R \xrightarrow{\cdot x} R \xrightarrow{\pi} R/(x) \longrightarrow 0 \, .$$

The first map is multiplication by x, and the second map is the canonical projection.

Example 0.10. Given an ideal I in a ring R, the inclusion map $\iota : I \to R$ and the canonical projection $\pi : R \to R/I$ give us the following short exact sequence:

$$0 \longrightarrow I \xrightarrow{\iota} R \xrightarrow{\pi} R/I \longrightarrow 0.$$

Example 0.11. Let $R = k[x]/(x^2)$. The following complex is exact:

$$\cdots \longrightarrow R \xrightarrow{\cdot x} R \xrightarrow{\cdot x} R \longrightarrow \cdots \, .$$

Indeed, the image and the kernel of multiplication by x are both (x).

Sometimes we can show that certain modules vanish or compute them explicitly when they do not vanish by seeing that they fit in some naturally constructed exact sequence involving other modules we understand better. We will discuss this in more detail when we talk about long exact sequences.

Remark 0.12. The complex $0 \longrightarrow M \xrightarrow{f} N \longrightarrow 0$ is exact if and only if f is an isomorphism.

Remark 0.13. The complex $0 \longrightarrow M \longrightarrow 0$ is exact if and only if $M = 0$.

Historically, chain complexes first appeared in topology. To study a topological space, one constructs a particular chain complex that arises naturally from information from the space, and then calculates its homology, which ends up encoding important topological information in the form of a sequence of abelian groups.

Definition 0.14 (Homology)**.** The **homology** of the complex $(C_\bullet, \partial_\bullet)$ is the sequence of R-modules

$$\mathrm{H}_n(C_\bullet) = \mathrm{H}_n(C) := \frac{\ker \partial_n}{\operatorname{im} \partial_{n+1}}.$$

The n**th homology** of $(C_\bullet, \partial_\bullet)$ is $\mathrm{H}_n(C)$. The submodules $Z_n(C_\bullet) = Z_n(C) := \ker \partial_n \subseteq C_n$ are called **cycles**, while the submodules $B_n(C_\bullet) = B_n(C) := \operatorname{im} \partial_{n+1} \subseteq C_n$ are called **boundaries**. One sometimes uses the word boundary to refer an element of $B_n(C)$ (an n-boundary), and the word cycle to refer to an element of $Z_n(C)$ (an n-cycle).

The homology of a complex measures how far our complex is from being exact at each point. Again, we can talk about the **cohomology** of a cochain complex instead, which we write as $\mathrm{H}^n(C)$; we will for now not worry about the distinction.

Remark 0.15. Note that $(C_\bullet, \partial_\bullet)$ is exact at n if and only if $\mathrm{H}_n(C_\bullet) = 0$.

Example 0.16. Let $R = k[x]/(x^3)$. Consider the following complex:

$$F_\bullet = \cdots \longrightarrow R \xrightarrow{\cdot x^2} R \xrightarrow{\cdot x^2} R \longrightarrow \cdots .$$

The image of multiplication by x^2 is (x^2), while the the kernel of multiplication by x^2 is $(x) \supseteq (x^2)$. For all n,

$$\mathrm{H}_n(F_\bullet) = (x)/(x^2) \cong R/(x).$$

Example 0.17. Let $\mathbb{Z} \xrightarrow{\pi} \mathbb{Z}/2\mathbb{Z}$ be the canonical projection map. Then

$$\underset{2}{C} = \mathbb{Z} \xrightarrow{4} \underset{1}{\mathbb{Z}} \xrightarrow{\pi} \underset{0}{\mathbb{Z}/2\mathbb{Z}}$$

is a complex of abelian groups, since the image of multiplication by 4 is $4\mathbb{Z}$, and that is certainly contained in $\ker \pi = 2\mathbb{Z}$. The homology of C is

$$\begin{aligned}
\mathrm{H}_n(C) &= 0 && \text{for } n \geqslant 3 \\
\mathrm{H}_2(C) &= \frac{\ker(\mathbb{Z} \xrightarrow{4} \mathbb{Z})}{\mathrm{im}(0 \longrightarrow \mathbb{Z})} = \frac{0}{0} = 0 \\
\mathrm{H}_1(C) &= \frac{\ker(\mathbb{Z} \xrightarrow{\pi} \mathbb{Z}/2\mathbb{Z})}{\mathrm{im}(\mathbb{Z} \xrightarrow{4} \mathbb{Z})} = \frac{2\mathbb{Z}}{4\mathbb{Z}} \cong \mathbb{Z}/2\mathbb{Z} \\
\mathrm{H}_0(C) &= \frac{\ker(\mathbb{Z}/2\mathbb{Z} \longrightarrow 0)}{\mathrm{im}(\mathbb{Z} \longrightarrow \mathbb{Z}/2\mathbb{Z})} = \frac{\mathbb{Z}/2\mathbb{Z}}{\mathbb{Z}/2\mathbb{Z}} = 0 \\
\mathrm{H}_n(C) &= 0 && \text{for } n < 0
\end{aligned}$$

Notice that our complex is exact at 2 and 0. The exactness at 2 says that the map $\mathbb{Z} \xrightarrow{4} \mathbb{Z}$ is injective, while exactness at 0 says that π is surjective.

Before we can continue any further into the world of homological algebra, we will need some categorical language. We will take a short break to introduce category theory, and then armed with that knowledge we will be ready to study homological algebra.

Chapter 1

Categories for the working homological algebraist

Most fields in modern mathematics follow the same basic recipe: there is a main type of object one wants to study – groups, rings, modules, topological spaces, etc – and a natural notion of arrows between these – group homomorphisms, ring homomorphisms, module homomorphisms, continuous maps, etc. The objects are often sets with some extra structure, and the arrows are often maps between the objects that preserve whatever that extra structure is. Category theory is born of this realization, by abstracting the basic notions that make math and studying them all at the same time. How many times have we felt a sense of déjà vu when learning about a new field of math? Category theory unifies all those ideas we have seen over and over in different contexts.

Category theory is an entire field of mathematics in its own right. As such, there is a lot to say about category theory, and unfortunately it doesn't all fit in the little time we have to cover it in this course. You are strongly encouraged to learn more about category theory, for example from [ML98] or [Rie17].

Before we go any further, note that there is a long and fun story about why we use the word *collection* when describing the objects in a category. Not all collections are allowed to be sets, an issue that was first discovered by Russel with his famous Russel's Paradox.[1] Russel exposed the fact that one has to be careful with how we formalize set theory. We follow the ZFC (Zermelo–Fraenkel with choice, short for the Zermelo–Fraenkel axioms plus the Axiom of Choice) axiomatization of set theory, and while we will not discuss the details of this formalization here, you are encouraged to read more on the subject.

1.1 Categories

A category consists of a collection of objects and arrows or morphisms between those objects. While these are often sets and some kind of functions between them, beware that this will not always be the case. We will use the words morphism and arrows interchangeably, though *arrow* has the advantage of reminding us we are not necessarily talking about functions.

[1] The collection of all sets that don't contain themselves cannot be a set. Do you see why?

Definition 1.1. A **category** $\mathscr{C}$ consists of three different pieces of data:

- a collection of **objects**, $\mathbf{ob}(\mathscr{C})$,
- for each two objects, say A and B, a collection $\mathrm{Hom}_{\mathscr{C}}(A, B)$ of **arrows** or **morphisms** from A to B, and
- for each three objects A, B, and C, a composition

$$\begin{aligned} \mathrm{Hom}_{\mathscr{C}}(A,B) \times \mathrm{Hom}_{\mathscr{C}}(B,C) &\longrightarrow \mathrm{Hom}_{\mathscr{C}}(A,C)\ . \\ (f,g) &\longmapsto g \circ f \end{aligned}$$

  We will often drop the $\circ$ and write simply gf for $g \circ f$.

These ingredients satisfy the following axioms:

1) The $\mathrm{Hom}_{\mathscr{C}}(A, B)$ are all disjoint. In particular, if f is an arrow in $\mathscr{C}$, we can talk about its **source** A and its **target** B as the objects such that $f \in \mathrm{Hom}_{\mathscr{C}}(A, B)$.
2) For each object A, there is an **identity arrow** $1_A \in \mathrm{Hom}_{\mathscr{C}}(A, A)$ such that $1_A \circ f = f$ and $g \circ 1_A = g$ for all $f \in \mathrm{Hom}_{\mathscr{C}}(B, A)$ and all $g \in \mathrm{Hom}_{\mathscr{C}}(A, B)$.
3) Composition is **associative**: $f \circ (g \circ h) = (f \circ g) \circ h$ for all appropriately chosen arrows.

Notation 1.2. We sometimes write $f\colon A \to B$ or $A \xrightarrow{f} B$ for an arrow $f \in \mathrm{Hom}(A, B)$.

Exercise 1.3. Prove that every element in a category has a unique identity morphism.

Here are some categories you have likely encountered before:

Example 1.4.

1) The category **Set** with objects all sets and arrows all functions between sets.
2) The category **Grp** whose objects are the collection of all groups, and whose arrows are all the homomorphisms of groups. The identity arrows are the identity homomorphisms.
3) The category **Ab** with objects all abelian groups, and arrows the homomorphisms of abelian groups. The identity arrows are the identity homomorphisms.
4) The category **Ring** of rings and ring homomorphisms. Contrary to what you may expect, this is not nearly as important as the next one.
5) The category R-**mod** of left modules over a fixed ring R and with R-module homomorphisms. Sometimes one writes R-**Mod** for this category, and reserve R-**mod** for the category of finitely generated R-modules with R-module homomorphisms. When $R = k$ is a field, the objects in the category k-**Mod** are k-vector spaces, and the arrows are linear transformations; we may instead refer to this category as **Vect-**k.

6) The category **Top** of topological spaces and continuous functions.

One may consider many variations of the categories above. Here are some variations on vector spaces:

Example 1.5. Let k be a field.

1) The collection of finite dimensional k-vector spaces with all linear transformations is a category.

2) The collection of all n-dimensional k-vector spaces with all linear transformations is a category.

3) The collection of all k-vector spaces (or n-dimensional vector spaces) with linear isomorphisms is a category.

4) The collection of all k-vector spaces (or n-dimensional vector spaces) with nonzero linear transformations is not a category, since it is not closed under composition.

5) The collection of all n-dimensional vector spaces with linear transformations of determinant 0 is not a category, since it does not have identity maps.

Here is an important variation of **Set**:

Example 1.6. The category $\mathbf{Set}^*$ of pointed sets has objects all pairs (X, x) of sets X and points $x \in X$, and for two pointed sets (X, x) and (Y, y), the morphisms from (X, x) to (Y, y) are functions $f\colon X \to Y$ such that $f(x) = y$, with the usual composition of functions.

Example 1.7. The empty category has no objects and no arrows.

While the collections of objects and arrows might not actually be sets, sometimes they are.

Definition 1.8. A category $\mathscr{C}$ is **locally small** if for all objects A and B in $\mathscr{C}$, $\mathrm{Hom}_{\mathscr{C}}(A, B)$ is a set. A category $\mathscr{C}$ is **small** if it is locally small and the collection of all objects in $\mathscr{C}$ is a set.

In fact, one can define a small category as one where the collection of all arrows is a set. It follows immediately that the collection of all objects is also a set, since it must be a subset of the set of arrows – for each object, there is an identity arrow.

Many important categories are at least locally small. For example, **Set** is locally small but not small. In a locally small category, we can now refer to its Hom-sets.

Categories where the objects are sets with some extra structure and the arrows are some kind of functions between the objects are called **concrete**. Not all categories are concrete.

Example 1.9. Given a partially ordered set $(X, \leqslant)$, we can regard X itself as a category: the objects are the elements of X, and for each x and y in X, $\mathrm{Hom}_X(x, y)$ is either a singleton if $x \leqslant y$ or empty if $x \not\leqslant y$. There is only one possible way to define composition, and the transitive property of $\leqslant$ guarantees that the composition of arrows is indeed well-defined: if

there is an arrow $i \to j$ and an arrow $j \to k$, then $i \leqslant j$ and $j \leqslant k$, so $i \leqslant k$ and thus there is a unique arrow $i \to k$. This category is clearly locally small, since all nonempty Hom-sets are in fact singletons. It is in fact small, since the objects are by construction the set X. We will denote this poset category by $\mathbf{PO}(X)$.

Example 1.10. For each positive integer n, the category $\mathbf{n}$ has n objects $0, 1, \ldots, n-1$ and $\mathrm{Hom}(i, j)$ is either empty if $i > j$ or a singleton if $i \leqslant j$. As Theorem 1.9, composition is defined in the only way possible, and things work out. This is the poset category for the poset $(\{0, 1, \ldots, n-1\}, \leqslant)$ with the usual $\leqslant$.

Example 1.11. Fix a field k. We define a category **Mat-k** with objects all positive integers, and given two positive integers a and b, the Hom-set $\mathrm{Hom}(a, b)$ consists of all $b \times a$ matrices with entries in k. The composition rule is given by product of matrices: given $A \in \mathrm{Hom}(a, b)$ and $B \in \mathrm{Hom}(b, c)$, the composition $B \circ A$ is the matrix $BA \in \mathrm{Hom}(a, c)$. For each object a, its identity arrow is given by the $a \times a$ identity matrix.

Example 1.12. Let G be a directed graph. We can construct a category from G as follows: the objects are the vertices of G, and the arrows are directed paths in the graph G. In this category, composition of arrows corresponds to concatenation of paths. For each object A, the identity arrow corresponds to the empty path from A to A.

Remark 1.13. A locally small category with just one element is completely determined by its unique Hom-set; it thus consists of a set S with an associative operation that has an identity element, which in this class is what we call a **semigroup**.[2]

A key insight we get from category theory is that many important concepts can be understood through diagrams. Homological algebra is in many ways the study of commutative diagrams. One way to formalize what a diagram is involves talking about functors, which we will discuss in Section 1.2; here is a more down to earth definition.

Definition 1.14. A **diagram** in a category $\mathscr{C}$ is a directed multigraph whose vertices are objects in C and whose arrows/edges are morphisms in $\mathscr{C}$. A commutative diagram in $\mathscr{C}$ is a diagram in which for each pair of vertices A and B, any two paths from A to B compose to the same morphism.

Example 1.15. The diagram

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle u}\downarrow & & \downarrow{\scriptstyle g} \\ C & \xrightarrow[v]{} & D \end{array}$$

commutes if and only if $gf = vu$.

There are some special types of arrows we will want to consider.

Definition 1.16. Let $\mathscr{C}$ be any category.

[2]Some authors prefer the term monoid.

- An arrow $f \in \mathrm{Hom}_{\mathscr{C}}(A, B)$ is **left invertible** if there exists $g \in \mathrm{Hom}_{\mathscr{C}}(B, A)$ such that $gf = 1_A$. In this case, we say that g is the **left inverse** of f. So g is a left inverse of f if the diagram

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ & \searrow^{1_A} & \downarrow{\scriptstyle g} \\ & & A \end{array}$$

  commutes.

- An arrow $f \in \mathrm{Hom}_{\mathscr{C}}(A, B)$ is **right invertible** if there exists $g \in \mathrm{Hom}_{\mathscr{C}}(B, A)$ such that $fg = 1_B$. In this case, we say that g is the **right inverse** of f. So g is a right inverse of f if the diagram

$$\begin{array}{ccc} B & \xrightarrow{g} & A \\ & \searrow^{1_B} & \downarrow{\scriptstyle f} \\ & & B \end{array}$$

  commutes.

- An arrow $f \in \mathrm{Hom}_{\mathscr{C}}(A, B)$ is an **isomorphism** if there exists $g \in \mathrm{Hom}_{\mathscr{C}}(B, A)$ such that $gf = 1_A$ and $fg = 1_B$. Unsurprisingly, such an arrow g is called the **inverse** of f. We say two objects A and B are **isomorphic** if there exists an isomorphism $A \to B$.

- An arrow $f \in \mathrm{Hom}(B, C)$ is **monic**, a **monomorphism**, or a **mono** if for all arrows

$$A \underset{g_2}{\overset{g_1}{\rightrightarrows}} B \xrightarrow{f} C$$

  if $fg_1 = fg_2$ then $g_1 = g_2$.

- Similarly, an arrow $f \in \mathrm{Hom}(A, B)$ is an **epi** or an **epimorphism** if for all arrows

$$A \xrightarrow{f} B \underset{g_2}{\overset{g_1}{\rightrightarrows}} C$$

  if $g_1 f = g_2 f$ then $g_1 = g_2$.

Here are some examples:

Exercise 1.17. Show that in **Set**, the monos coincide with the injective functions and the epis coincide with the surjective functions.

Example 1.18.

1) In **Grp**, **Ring**, and R-**Mod** the isomorphisms are the morphisms that are bijective functions.

2) In contrast, in **Top** the isomorphisms are the homeomorphisms, which are the bijective continuous functions with continuous inverses. These are *not* the same thing as just the bijective continuous functions.

Exercise 1.19. Show that in any category, every isomorphism is both epi and mono.

Exercise 1.20. Show that the usual inclusion $\mathbb{Z} \longrightarrow \mathbb{Q}$ is an epi in the category **Ring**.

This *should* feel weird: it says being epi and being surjective are *not* the same thing. Similarly, being monic and being injective are *not* the same thing.

Exercise 1.21. Show that the canonical projection $\mathbb{Q} \longrightarrow \mathbb{Q}/\mathbb{Z}$ is a mono in the category of divisible abelian groups.[3]

Exercise 1.22. Show that given any poset P, in the poset category of P every morphism is both monic and epic, but no nonidentity morphism has a left or right inverse.

There are some special types of objects we will want to consider.

Definition 1.23. Let $\mathscr{C}$ be a category. An **initial object** in $\mathscr{C}$ is an object i such that for every object x in $\mathscr{C}$, $\mathrm{Hom}_{\mathscr{C}}(i, x)$ is a singleton, meaning there exists a unique arrow $i \longrightarrow x$. A **terminal object** in $\mathscr{C}$ is an object t such that for every object x in $\mathscr{C}$, $\mathrm{Hom}_{\mathscr{C}}(x, t)$ is a singleton, meaning there exists a unique arrow $x \longrightarrow t$. A **zero object** in $\mathscr{C}$ is an object that is both initial and terminal.

Exercise 1.24. Initial objects are unique up to unique isomorphism. Terminal objects are unique up to unique isomorphism.

So we can talk about *the* initial object, *the* terminal object, and *the* zero object, if they exist.

Example 1.25.

a) The empty set is initial in **Set**. Any singleton is terminal. Since the empty set and a singleton are not isomorphic in **Set**, there is no zero object in **Set**.

b) The 0 module is the zero object in R-**Mod**.

c) The trivial group $\{e\}$ is the zero object in **Grp**.

d) In the category of rings, $\mathbb{Z}$ is the initial object, but there is no terminal object unless we allow the 0 ring.

e) There are no initial nor terminal objects in the category of fields.

We will now continue to follow a familiar pattern and define the related concepts one can guess should be defined.

Definition 1.26. A **subcategory** $\mathscr{C}$ of a category $\mathscr{D}$ consists of a subcollection of the objects of $\mathscr{D}$ and a subcollection of the morphisms of $\mathscr{D}$ such that the following hold:

- For every object C in $\mathscr{C}$, the arrow $1_C \in \mathrm{Hom}_{\mathscr{D}}(C, C)$ is an arrow in $\mathscr{C}$.

[3] An abelian group A is divisible if for every $a \in A$ and every positive integer n there exists $b \in A$ such that $nb = a$.

- For every arrow in $\mathscr{C}$, its source and target in $\mathscr{D}$ are objects in $\mathscr{C}$.
- For every pair of arrows f and g in $\mathscr{C}$ such that fg is an arrow that makes sense in $\mathscr{D}$, fg is an arrow in $\mathscr{C}$.

In particular, $\mathscr{C}$ is a category in its own right.

Example 1.27. The category of finitely generated R-modules with R-module homomorphisms is a subcategory of R-**Mod**.

Definition 1.28. A subcategory $\mathscr{C}$ of $\mathscr{D}$ is a **full subcategory** if $\mathscr{C}$ includes *all* of the arrows in $\mathscr{D}$ between any two objects in $\mathscr{C}$.

Example 1.29.

a) The category **Ab** of abelian groups is a full subcategory of **Grp**.

b) Since every group is a set, and every homomorphism is a function, **Grp** is a subcategory of **Set**. However, not every function between two groups is a group homomorphism, so **Grp** is not a full subcategory of **Set**.

c) The category whose objects are all sets and with arrows all bijections is a subcategory of **Set** that is not full.

Here is another way of constructing a new category out of an old one.

Definition 1.30. Let $\mathscr{C}$ be a category. The **opposite category** of $\mathscr{C}$, denoted $\mathscr{C}^{\mathrm{op}}$, is a category whose objects are the objects of $\mathscr{C}$, and such that each arrow $f \in \mathrm{Hom}_{\mathscr{C}^{\mathrm{op}}}(A, B)$ is the same as some arrow in $\mathrm{Hom}_{\mathscr{C}}(B, A)$. The composition fg of two morphisms f and g in $\mathscr{C}^{\mathrm{op}}$ is defined as the composition gf in $\mathscr{C}$.

Many objects and concepts one might want to describe are obtained from existing ones by flipping the arrows. Opposite categories give us the formal framework to talk about such things. We will often want to refer to **dual** notions, which will essentially mean considering the same notion in a category $\mathscr{C}$ and in the opposite category $\mathscr{C}^{\mathrm{op}}$; in practice, this means we should flip all the arrows involved. We will see examples of this later on.

The dual category construction gives us a formal framework to talk about **dual notions**. We will often make a statement in a category $\mathscr{C}$ and make comments about the **dual statement**; in practice, this corresponds to simply switching the way all arrows go. Here are some examples of dual notions and statements:

source	target
epi	mono
g is a right inverse for f	g is a left inverse for f
f is invertible	f is invertible
initial objects	terminal objects
homology	cohomology

The prefix co- is often used to denote the dual of something, such as in *co*homology. Note that the dual of the dual is the original statement; formally, $(\mathscr{C}^{\mathrm{op}})^{\mathrm{op}} = \mathscr{C}$. Sometimes we can easily prove a statement by dualizing; however, this is not always straightforward, and one needs to carefully dualize all portions of the statement in question. Nevertheless, Sanders MacLane, one of the fathers of category theory, wrote that "If any statement about a category is deducible from the axioms for a category, the dual statement is likely deducible" [Mac50]. One of the upshots of duality is that any theorem in category theory must simultaneously prove two theorems: the original statement and its dual. But for this to hold, we need proofs that use the abstraction of a purely categorical proof.

Opposite categories are more interesting than they might appear at first; there is more than just flipping all the arrows. For example, consider the opposite category of **Set**. For any nonempty set X, there is a unique morphism in **Set** (a function) $i : \emptyset \to X$, but there are no functions $X \to \emptyset$, so $i^{\mathrm{op}} \colon \emptyset \to X$ is not a function. Thus thinking about $\mathbf{Set}^{\mathbf{op}}$ is a bit difficult. One can show that this is the category of complete atomic Boolean algebras – but we won't concern ourselves with what that means.

1.2 Functors

Many mathematical constructions are *functorial*, in the sense that they behave well with respect to morphisms. In the formalism of category theory, this means that we can think of a functorial construction as a functor.

Definition 1.31. Let $\mathscr{C}$ and $\mathscr{D}$ be categories. A **covariant functor** $F \colon \mathscr{C} \longrightarrow \mathscr{D}$ is a mapping that assigns to each object A in $\mathscr{C}$ an object $F(A)$ in $\mathscr{D}$, and to each arrow $f \in \mathrm{Hom}_{\mathscr{C}}(A, B)$ an arrow $F(f) \in \mathrm{Hom}_{\mathscr{D}}(F(A), F(B))$, such that

- F preserves composition: $F(fg) = F(f)F(g)$ for all composable arrows f and g in $\mathscr{C}$.
- F preserves the identity arrows: $F(1_A) = 1_{F(A)}$ for all objects A in $\mathscr{C}$.

A **contravariant functor** $F \colon \mathscr{C} \longrightarrow \mathscr{D}$ is a mapping that assigns to each object A in $\mathscr{C}$ an object $F(A)$ in $\mathscr{D}$, and to each arrow $f \in \mathrm{Hom}_{\mathscr{C}}(A, B)$ an arrow $F(f) \in \mathrm{Hom}_{\mathscr{D}}(F(B), F(A))$, such that

- F preserves composition: $F(fg) = F(g)F(f)$ for all composable arrows f and g in $\mathscr{C}$.
- F preserves the identity arrows: $F(1_A) = 1_{F(A)}$ for all objects A in $\mathscr{C}$.

So a contravariant functor is a functor that flips all the arrows. We can also describe a contravariant functor as a covariant functor from $\mathscr{C}$ to the opposite category of $\mathscr{D}$, $\mathscr{D}^{\mathrm{op}}$.

Remark 1.32. A contravariant functor $F \colon \mathscr{C} \longrightarrow \mathscr{D}$ can be thought of as a covariant functor $\mathscr{C}^{\mathrm{op}} \longrightarrow \mathscr{D}$, or also as a covariant functor $\mathscr{C} \longrightarrow \mathscr{D}^{\mathrm{op}}$. If using one of these conventions, one needs to be careful, however, when composing functors, so that the respective sources and targets match up correctly. While we haven't specially discussed how one composes functors, it should be clear that applying a functor $F \colon \mathscr{C} \longrightarrow \mathscr{D}$ and $G \colon \mathscr{D} \longrightarrow \mathscr{E}$ is the same as applying a functor $\mathscr{C} \longrightarrow \mathscr{E}$, which we can write as GF.

For example, if $F : \mathscr{C} \longrightarrow \mathscr{D}$ and $G : \mathscr{D} \longrightarrow \mathscr{E}$ are both contravariant functors, the composition $GF : \mathscr{C} \longrightarrow \mathscr{E}$ is a covariant functor, since

$$\begin{array}{ccc} A & & F(A) & & GF(A) \\ {\scriptstyle f}\downarrow & \rightsquigarrow & {\scriptstyle F(f)}\uparrow & \rightsquigarrow & {\scriptstyle GF(f)}\downarrow \\ B & & F(B) & & GF(B) \end{array}$$

So we could think of F as a covariant functor $\mathscr{C} \longrightarrow \mathscr{D}^{\mathrm{op}}$ and G as a covariant functor $\mathscr{D}^{\mathrm{op}} \longrightarrow \mathscr{E}$. Similarly, if $F : \mathscr{C} \longrightarrow \mathscr{D}$ is a covariant functor and $G : \mathscr{D} \longrightarrow \mathscr{E}$ is a contravariant functor, $GF : \mathscr{C} \longrightarrow \mathscr{E}$ is a contravariant functor. In this case, we can think of G as a covariant functor $\mathscr{D} \longrightarrow \mathscr{E}^{\mathrm{op}}$, so that GF is now a covariant functor $\mathscr{C} \longrightarrow \mathscr{E}^{\mathrm{op}}$.

Exercise 1.33. Show that functors preserve isomorphisms.

Remark 1.34. Any functor sends isos to isos, since it preserves compositions and identities.

Example 1.35. Here are some examples of functors you may have encountered before.

a) Many categories one may think about are concrete categories, where the objects are sets with some extra structure, and the arrows are functions between those sets that preserved that extra structure. The **forgetful functor** from such a category to **Set** is the functor that, just as the name says, *forgets* that extra structure, and sees only the underlying sets and functions of sets. For example, the forgetful functor **Gr** $\to$ **Set** sends each group to its underlying set, and each group homomorphism to the corresponding function of sets.

b) The identity functor $1_{\mathscr{C}}$ on any category $\mathscr{C}$ does what the name suggests: it sends each object to itself and each arrow to itself.

c) Given an object C in a category $\mathscr{C}$, the functor $\Delta C : \mathscr{C} \to \mathscr{C}$ that sends every object to C and every arrow to 1_C is called the **constant functor** at C.

d) Given a group G, the subgroup $[G, G]$ of G generated by the set of commutators

$$\{ghg^{-1}h^{-1} \mid g, h \in G\}$$

is a normal subgroup, and the quotient $G^{\mathrm{ab}} := G/[G, G]$ is the **abelianization** of G. The group G^{ab} is abelian. Given a group homomorphism $f : G \to H$, f automatically takes commutators to commutators, so it induces a homomorphism $\tilde{f} : G^{\mathrm{ab}} \to H^{\mathrm{ab}}$. More precisely, abelianization gives a covariant functor from **Grp** to **Ab**.

e) The unit group functor $-^* :$ **Ring** $\to$ **Grp** sends a ring R to its group of units R^*. To see this is indeed a functor, we should check it behaves well on morphisms; and indeed if $f : R \to S$ is a ring homomorphism, and $u \in R^*$ is a unit in R, then

$$f(u)f(u^{-1}) = f(uu^{-1}) = f(1_R) = 1_S,$$

so $f(u)$ is a unit in S. Thus f induces a function $R^* \to S^*$ given by restriction of f to R^*, which must therefore be a group homomorphism since f preserves products.

f) Fix a field k. Given a vector space V, the set V^* of linear transformations from V to k is a k-vector space, the **dual vector space** of V. If $\varphi : W \to V$ is a linear transformation and $\ell \colon V \to k$ is an element of V^*, then $\ell \circ \varphi \colon W \to k$ is in W^*. Doing this for all elements $\ell \in V^*$ gives a function $\varphi^* : V^* \to W^*$, and one can show that φ^* is a linear transformation. The assignment that sends each vector space V to its dual vector space V^* and each linear transformation φ to φ^* is a contravariant functor **Vect-**$k \to$ **Vect-**k.

g) Localization is a functor. Let R be a ring and W be a multiplicatively closed set in R. The localization at W induces a a functor R**-mod** $\longrightarrow W^{-1}R$**-mod**: this functor sends each R-module M to $W^{-1}M$, and each R-module homomorphism $\alpha \colon M \to N$ to the R-module homomorphism $W^{-1}\alpha : W^{-1}M \to W^{-1}N$.

Remark 1.36. If we apply a covariant functor to a diagram, then we get a diagram of the same shape:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle u}\downarrow & & \downarrow{\scriptstyle g} \\ C & \xrightarrow[v]{} & D \end{array} \qquad \overset{F}{\rightsquigarrow} \qquad \begin{array}{ccc} F(A) & \xrightarrow{F(f)} & F(B) \\ {\scriptstyle F(u)}\downarrow & & \downarrow{\scriptstyle F(g)} \\ F(C) & \xrightarrow[F(v)]{} & F(D) \end{array}$$

However, if we apply a contravariant functor to the same diagram, we get a similar diagram but with the arrows reversed:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle u}\downarrow & & \downarrow{\scriptstyle g} \\ C & \xrightarrow[v]{} & D \end{array} \qquad \overset{F}{\rightsquigarrow} \qquad \begin{array}{ccc} F(A) & \xleftarrow{F(f)} & F(B) \\ {\scriptstyle F(u)}\uparrow & & \uparrow{\scriptstyle F(g)} \\ F(C) & \xleftarrow[F(v)]{} & F(D) \end{array}$$

Definition 1.37. The category **Cat** has objects all small categories and arrows all functors between them.

If we think about functors as functions between categories, it's natural to consider what would be the appropriate versions of the notions of injective or surjective.

Definition 1.38. A covariant functor $F : \mathscr{C} \longrightarrow \mathscr{D}$ between locally small categories is

- **faithful** if all the functions of sets

$$\begin{aligned} \mathrm{Hom}_{\mathscr{C}}(A, B) &\longrightarrow \mathrm{Hom}_{\mathscr{D}}(F(A), F(B)) \\ f &\longmapsto F(f) \end{aligned}$$

  are injective.

- **full** if all the functions of sets

$$\begin{aligned} \mathrm{Hom}_{\mathscr{C}}(A, B) &\longrightarrow \mathrm{Hom}_{\mathscr{D}}(F(A), F(B)) \\ f &\longmapsto F(f) \end{aligned}$$

  are surjective.

- **fully faithful** if it is full and faithful.
- **essentially surjective** if every object d in $\mathscr{D}$ is isomorphic to $F(c)$ for some c in $\mathscr{C}$.
- **essentially injective** if for all objects x and y in $\mathscr{C}$, if $F(x)$ is isomorphic to $F(y)$ in $\mathscr{D}$ then x is isomorphic to d in $\mathscr{C}$.
- an **embedding** if it is fully faithful and injective on objects, meaning that for all objects x and y in $\mathscr{C}$, if $F(x) = F(y)$ then $x = y$.

Example 1.39. The forgetful functor R-**Mod** $\longrightarrow$ **Set** is faithful since any two maps of R-modules with the same source and target coincide if and only if they are the same function of sets. This functor is not full, since not every function between the underlying sets of two R-modules is an R-module homomorphism.

Remark 1.40. A fully faithful functor is not necessarily injective on objects, but it is injective on objects up to isomorphism.

Remark 1.41. A subcategory $\mathscr{C}$ of $\mathscr{D}$ is full if the inclusion functor $\mathscr{C} \longrightarrow \mathscr{D}$ is full.

Exercise 1.42. Show that every fully faithful functor $F\colon \mathscr{C} \to \mathscr{D}$ reflects isos:

a) If f is an arrow in $\mathscr{C}$ such that $F(f)$ is an iso, then f is an iso.

b) If $F(X)$ and $F(Y)$ are isomorphic, then the objects X and Y are isomorphic in $\mathscr{C}$.

Note that the converses of these statements hold for any functor.

To close this section, here are the two of the most important functors we will discuss this semester:

Definition 1.43. Let $\mathscr{C}$ be a locally small category. An object A in $\mathscr{C}$ induces two Hom functors:

- The covariant functor $\mathrm{Hom}_{\mathscr{C}}(A, -) : \mathscr{C} \longrightarrow$ **Set** is defined as follows:

$$\begin{array}{lcc} & \mathscr{C} & \longrightarrow \mathbf{Set} \\ \text{on objects:} & X & \longmapsto \mathrm{Hom}_{\mathscr{C}}(A, X) \\ \text{on arrows:} & \begin{array}{c} B \\ {\scriptstyle f}\downarrow \\ C \end{array} & \rightsquigarrow \quad \begin{array}{c} \mathrm{Hom}_{\mathscr{C}}(A, B) \\ \downarrow \\ \mathrm{Hom}_{\mathscr{C}}(A, C) \end{array} \quad \begin{array}{c} \ni g \\ \downarrow \\ \ni f \circ g \end{array} \end{array}$$

  We read $\mathrm{Hom}_{\mathscr{C}}(A, -)$ as *Hom from* A, and may refer to this functor as the covariant functor **represented by** A. Given an arrow f in $\mathscr{C}$, we write $f_* := \mathrm{Hom}_{\mathscr{C}}(A, f)$. It is easier to see what f_* does through the following commutative diagram:

$$f_* = \mathrm{Hom}_{\mathscr{C}}(A, f): \qquad \begin{array}{ccc} A & \xrightarrow{\ g\ } & B \\ & {\scriptstyle f_*(g)=fg}\searrow & \downarrow {\scriptstyle f} \\ & & C \end{array}$$

- The contravariant functor $\mathrm{Hom}_{\mathscr{C}}(-,B):\mathscr{C}\longrightarrow \mathbf{Set}$ is defined as follows:

$$\begin{array}{lccc} & \mathscr{C} & \longrightarrow & \mathbf{Set} \\ \text{on objects:} & X & \longmapsto & \mathrm{Hom}_{\mathscr{C}}(X,B) \\ \text{on arrows:} & \begin{array}{c} A \\ {\scriptstyle f}\downarrow \\ C \end{array} & \rightsquigarrow & \begin{array}{c} \mathrm{Hom}_{\mathscr{C}}(A,B) \\ \uparrow \\ \mathrm{Hom}_{\mathscr{C}}(C,B) \end{array} \qquad \begin{array}{c} \ni g\circ f \\ \uparrow \\ \ni g \end{array} \end{array}$$

We read $\mathrm{Hom}_{\mathscr{C}}(-,B)$ as *Hom to* B, and we may refer to this functor as the contravariant functor **represented by** B. Given an arrow f in $\mathscr{C}$, we write $f^*:=\mathrm{Hom}_{\mathscr{C}}(-,B)$. It is easier to see what f^* does through the following commutative diagram:

$$f^*=\mathrm{Hom}_{\mathscr{C}}(f,B):\qquad \begin{array}{ccc} A & \xrightarrow{\ f\ } & C \\ & {\scriptstyle f^*(g)=gf}\searrow & \downarrow{\scriptstyle g} \\ & & B \end{array}$$

Exercise 1.44. Check that $\mathrm{Hom}(A,-)$ and $\mathrm{Hom}(-,B)$ are indeed functors.

We will be particularly interested in the Hom-functors in the category R-**mod**, which we will study in detail in a later chapter.

1.3 Natural transformations

Definition 1.45. Let F and G be covariant functors $\mathscr{C}\longrightarrow\mathscr{D}$. A **natural transformation** between F and G is a mapping that to each object A in $\mathscr{C}$ assigns an arrow $\eta_A\in\mathrm{Hom}_{\mathscr{D}}(F(A),G(A))$ such that for all $f\in\mathrm{Hom}_{\mathscr{C}}(A,B)$, the diagram

$$\begin{array}{ccc} F(A) & \xrightarrow{\eta_A} & G(A) \\ {\scriptstyle F(f)}\downarrow & & \downarrow{\scriptstyle G(f)} \\ F(B) & \xrightarrow[\eta_B]{} & G(B) \end{array}$$

commutes. We sometimes write

$$\mathscr{C}\ \underset{G}{\overset{F}{\Downarrow\eta}}\ \mathscr{D}$$

or simply $\eta\colon F\implies G$ to indicate that η is a natural transformation from F to G.

Definition 1.46. Let F and G be contravariant functors $\mathscr{C} \longrightarrow \mathscr{D}$. A **natural transformation** between F and G is a mapping that to each object A in $\mathscr{C}$ assigns an arrow $\eta_A \in \mathrm{Hom}_{\mathscr{D}}(F(A), G(A))$ such that for all $f \in \mathrm{Hom}_{\mathscr{C}}(A, B)$, the diagram

$$\begin{array}{ccc} F(A) & \xrightarrow{\eta_A} & G(A) \\ {\scriptstyle F(f)}\uparrow & & \uparrow{\scriptstyle G(f)} \\ F(B) & \xrightarrow[\eta_B]{} & G(B) \end{array}$$

commutes.

Often, when studying a particular topic, we sometimes say a certain map is *natural* to mean that there is actually a natural transformation behind it.

Example 1.47. Recall the abelianization functor we discussed in Theorem 1.35. The abelianization comes equipped with a natural projection map $\pi_G : G \longrightarrow G^{\mathrm{ab}}$, the usual quotient map from G to a normal subgroup. Here we mean natural in two different ways: both that this is the common sense map to consider, and that this is in fact coming from a natural transformation. What's happening behind the scenes is that abelianization is a functor $\mathrm{ab}\colon \mathbf{Grp} \longrightarrow \mathbf{Grp}$. On objects, the abelianizations functor is defined as $G \mapsto G^{\mathrm{ab}}$. Given an arrow, meaning a group homomorphism $G \xrightarrow{f} H$, one can check that $[G, G]$ is contained in the kernel of $\pi_H f$, so $\pi_H f$ factors through G^{ab}, and there exists a group homomorphism f^{ab} making the following diagram commute:

$$\begin{array}{ccc} G & \xrightarrow{\pi_G} & G^{\mathrm{ab}} \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f^{\mathrm{ab}}} \\ H & \xrightarrow[\pi_H]{} & H^{\mathrm{ab}} \end{array} .$$

So the abelianization functor takes the arrow f to f^{ab}. The commutativity of the diagram above says that $\pi_{\_}$ is a natural transformation π between the identity functor on **Grp** and the abelianization functor, which we can write more compactly as

$$\mathbf{Grp} \underset{\mathrm{ab}}{\overset{\mathrm{id}}{\Downarrow \pi}} \mathbf{Grp} .$$

Example 1.48. The determinant gives rise to a natural transformation. Fix an integer $n \geqslant 1$, and consider the GL_n functor

$$\mathrm{GL}_n\colon \mathbf{Ring} \to \mathbf{Grp}$$

that takes each ring R to the group GL_n of invertible $n \times n$ matrices with entries in R, and that takes each ring homomorphism $f\colon R \to S$ to the map

$$\mathrm{GL}_n(f)\colon \mathrm{GL}_n(R) \to \mathrm{GL}_n(S)$$

that applies f to all the entries of each matrix $A \in \mathrm{GL}_n(R)$, and which can be shown to be a group homomorphism. We claim that the determinant is a natural transformation from GL_n to the unit functor $(-)^*$ we defined in Theorem 1.35. First, note that the determinant of an invertible matrix is a unit, so the determinant gives a map $\mathrm{GL}_n(R) \to R^*$. Moreover, given any ring homomorphism $f\colon R \to S$, we have a commutative diagram

$$\begin{array}{ccc} \mathrm{GL}_n(R) & \xrightarrow{\det} & R^* \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ \mathrm{GL}_n(S) & \xrightarrow[\det]{} & S^*. \end{array}$$

Above we identified f with both the map $\mathrm{GL}_n(f)$ obtained by applying f to all coordinates of A and the restriction of f to the unit groups, meaning the image of f under the units functor. This commutative diagram just encodes the fact that taking determinants commutes with applying f: for any invertible $n \times n$ matrix A,

$$f(\det(A)) = \det(f(A)).$$

Definition 1.49. A **natural isomorphism** is a natural transformation η where each η_A is an isomorphism.

Exercise 1.50. Show that a natural transformation $\eta\colon F \implies G$ is a natural isomorphism if and only if there exists a natural transformation $\mu : G \implies F$ such that $\eta \circ \mu$ is the identity natural isomorphism on G and $\mu \circ \eta$ is the identity natural isomorphism on F.

Warning: there are many theorems that say that a particular isomorphism is natural; however, not all isomorphisms are natural! Whenever S is an infinite set, the sets $S \times S$ are in bijection S, but no such bijection can be natural. Details below.

Exercise 1.51. Let $\mathbf{Set}^\infty$ be the full subcategory of $\mathbf{Set}$ consisting of all infinite sets. Let

$$F\colon \mathbf{Set}^\infty \to \mathbf{Set}^\infty$$

be the functor that on objects is given by the rule $F(S) = S \times S$, and on morphisms is given by $F(f) = (f, f)$. Show that there is no natural isomorphism $\eta : F \Rightarrow 1_{\mathbf{Set}^\infty}$.

Definition 1.52. Two categories $\mathscr{C}$ and $\mathscr{D}$ are **equivalent** if there exist functors $F\colon \mathscr{C} \to \mathscr{D}$ and $G\colon \mathscr{D} \to \mathscr{C}$ and two natural isomorphisms $\alpha\colon GF \implies 1_{\mathscr{C}}$ and $\beta\colon FG \implies 1_{\mathscr{D}}$. We say that a functor $F\colon \mathscr{C} \to \mathscr{D}$ is an **equivalence of categories** if there exists a functor G and natural isomorphisms α and β as above.

If one assumes the Axiom of Choice, this is the right notion of isomorphism of two categories (though not in the categorical sense!); better said, two categories that are equivalent are essentially the same. Note that this does not mean that there is a bijection between the objects of $\mathscr{C}$ and the objects of $\mathscr{D}$. In fact, one can show that a functor is an equivalence of categories if and only if it is fully faithful and essentially surjective – though this fact requires the Axiom of Choice!

Exercise 1.53. Let $\mathscr{C}$ be the category with one object C and a unique arrow 1_C. Let $\mathscr{D}$ be the category with two objects D_1 and D_2 and four arrows: the identities 1_{D_i} and two isomorphisms $\alpha : D_1 \to D_2$ and $\beta : D_2 \to D_1$. Let $\mathscr{E}$ be the category with two objects E_1 and E_2 and only two arrows, 1_{E_1} and 1_{E_2}.

a) Show that $\mathscr{C}$ and $\mathscr{D}$ are equivalent categories.

b) Show that $\mathscr{C}$ and $\mathscr{E}$ are not equivalent categories.

Definition 1.54. Let $F, G : \mathscr{C} \longrightarrow \mathscr{D}$ be two functors between the categories $\mathscr{C}$ and $\mathscr{D}$. We write $\mathrm{Nat}(F, G)$ for the collection of all natural transformations $F \implies G$. Given two categories $\mathscr{C}$ and $\mathscr{D}$, one can build a **functor category**[4] with objects all covariant functors $\mathscr{C} \longrightarrow \mathscr{D}$, and arrows the corresponding natural transformations. This category is denoted $\mathscr{D}^{\mathscr{C}}$. Sometimes one writes $\mathrm{Hom}(F, G)$ for $\mathrm{Nat}(F, G)$, but we will avoid that, as it might make things even more confusing.

For the functor category to truly be a category, though, we need to know how to compose natural transformations.

Remark 1.55. Consider natural transformations

$$\mathscr{C} \underset{G}{\overset{F}{\Downarrow \varphi}} \mathscr{D} \qquad \text{and} \qquad \mathscr{C} \underset{H}{\overset{G}{\Downarrow \eta}} \mathscr{D}.$$

We can compose them for form a new natural transformation

$$\mathscr{C} \underset{H}{\overset{F}{\Downarrow \eta\varphi}} \mathscr{D}.$$

For each object C in $\mathscr{C}$, $\eta\varphi$ sends C to the arrow

$$F(C) \xrightarrow{\varphi_C} G(C) \xrightarrow{\eta_C} H(C).$$

This makes the diagram

$$\begin{array}{ccccc} F(A) & \xrightarrow{\varphi_A} & G(A) & \xrightarrow{\eta_A} & H(A) \\ {\scriptstyle F(f)}\downarrow & & {\scriptstyle G(f)}\downarrow & & \downarrow{\scriptstyle H(f)} \\ F(B) & \xrightarrow[\varphi_B]{} & G(B) & \xrightarrow[\eta_B]{} & H(B) \end{array}$$

commute; replacing the horizontal arrows with the composition gives us the commutative diagram

$$\begin{array}{ccc} F(A) & \xrightarrow{\eta_A\varphi_A} & H(A) \\ {\scriptstyle F(f)}\downarrow & & \downarrow{\scriptstyle H(f)} \\ F(B) & \xrightarrow[\eta_B\varphi_B]{} & H(B) \end{array}$$

which encodes the fact that $\eta\varphi$ is a natural transformation.

[4] Yes, the madness is neverending.

The functors that are naturally isomorphic to some Hom functor are important.

Definition 1.56. A covariant functor $F : \mathscr{C} \longrightarrow \mathbf{Set}$ is **representable** if there exists an object A in $\mathscr{C}$ such that F is naturally isomorphic to $\mathrm{Hom}_{\mathscr{C}}(A, -)$. A contravariant functor $F : \mathscr{C} \longrightarrow \mathbf{Set}$ is **representable** if there exists an object B in $\mathscr{C}$ such that F is naturally isomorphic to $\mathrm{Hom}_{\mathscr{C}}(-, B)$.

Example 1.57. We claim that the identity functor $\mathbf{Set} \longrightarrow \mathbf{Set}$ is representable. Let $\mathbf{1}$ be a singleton set. Given any set X, there is a bijection between elements $x \in X$ and functions $\mathbf{1} \longrightarrow X$ sending the one element in $\mathbf{1}$ to each x. Moreover, given any other set Y, and a function $f : X \longrightarrow Y$, our bijections make the following diagram commute:

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{Set}}(\mathbf{1}, X) & \xrightarrow{\cong} & X \\ {\scriptstyle f_*}\downarrow & & \downarrow{\scriptstyle f} \\ \mathrm{Hom}_{\mathbf{Set}}(\mathbf{1}, Y) & \xrightarrow{\cong} & Y. \end{array}$$

This data gives a natural isomorphism between the identity functor and $\mathrm{Hom}_{\mathbf{Set}}(\mathbf{1}, -)$.

Exercise 1.58. Show that the forgetful functor $\mathbf{Grp} \longrightarrow \mathbf{Set}$ is representable.

Exercise 1.59. Given a ring R, show that the forgetful functor $R\text{-}\mathbf{mod} \longrightarrow \mathbf{Set}$ is representable.

The Yoneda Lemma tells us that in order to study a locally small category $\mathscr{C}$, it is in many ways sufficient to study the category of functors from $\mathscr{C}$ to $\mathbf{Set}$, and that representable functors are the most important functors of all.

1.4 The Yoneda Lemma

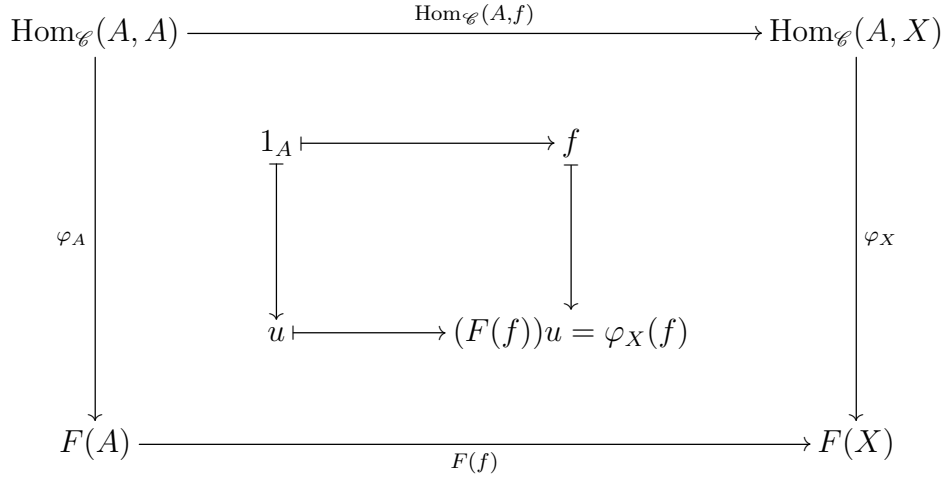
Even though this is only a short introduction to category theory, we would be remiss not to mention the Yoneda Lemma, arguably the most important statement in category theory.

Theorem 1.60 (Yoneda Lemma)**.** *Let $\mathscr{C}$ be a locally small category, and fix an object A in $\mathscr{C}$. Let $F\colon \mathscr{C} \longrightarrow \mathbf{Set}$ be a covariant functor. Then there is a bijection*

$$Nat(\mathrm{Hom}_{\mathscr{C}}(A,-), F) \xrightarrow{\gamma} F(A)\ .$$

Moreover, this correspondence is natural in both A and F.

Proof. Let φ be a natural transformation in $\mathrm{Nat}(\mathrm{Hom}_{\mathscr{C}}(A,-), F)$. The proof is essentially the following diagram:



Our bijection will be defined by

$$\gamma(\varphi) := \varphi_A(1_A).$$

We should first check that this makes sense: arrows in **Set** are just functions between sets, and so φ_A is a function of sets $\mathrm{Hom}_{\mathscr{C}}(A,A) \longrightarrow F(A)$. Also, $\mathrm{Hom}_{\mathscr{C}}(A,A)$ is a set that contains at least the element 1_A, and $\varphi_A(1_A)$ is some element in the set $F(A)$.

Given any fixed arrow $f \in \mathrm{Hom}_{\mathscr{C}}(A,X)$, the fact that φ is a natural transformation translates into the outer commutative diagram. In particular, the functions of sets $F(f)\varphi_A$ and $\varphi_X \mathrm{Hom}_{\mathscr{C}}(A,f)$ coincide, and must in particular take 1_A to the same element in $F(X)$. This is the commutativity of the inner diagram, with $u := \varphi_A(1_A)$.

The commutativity of the diagram above says that φ is completely determined by $\varphi_A(1_A)$, since for any other object X in $\mathscr{C}$ and any arrow $f \in \mathrm{Hom}_{\mathscr{C}}(A,X)$, we necessarily have $\varphi_X(f) = F(f)\varphi_A(1_A)$. Thus if φ and η are distinct natural transformations, then there exists some object X and some $f \in \mathrm{Hom}_{\mathscr{C}}(A,X)$ such that

$$\varphi_X(f) \neq \eta_X(f), \quad \text{so } F(f)\varphi_A(1_A) \neq F(f)\eta_A(1_A) \quad \text{and thus } \varphi_A(1_A) \neq \eta_A(1_A).$$

In particular, our map $\gamma(\varphi) = \varphi_A(1_A)$ is injective.

Moreover, note that each choice of $u \in F(A)$ gives rise to a different natural transformation φ by setting $\varphi_X(f) = F(f)u$. To check that this is in fact a natural transformation, one needs to check that for all arrows $g\colon X \to Y$, the diagram

$$\begin{array}{ccc} \mathrm{Hom}_{\mathscr{C}}(A,X) & \xrightarrow{\varphi_X} & F(X) \\ {\scriptstyle g_*}\downarrow & & \downarrow{\scriptstyle F(g)} \\ \mathrm{Hom}_{\mathscr{C}}(A,Y) & \xrightarrow[\varphi_Y]{} & F(Y) \end{array}$$

commutes. And indeed, given any $f \in \mathrm{Hom}_{\mathscr{C}}(A,X)$,

$$\begin{aligned} F(g)\circ\varphi_X(f) &= F(g)F(f)u && \text{by definition of } \varphi \\ &= F(gf)u && \text{since } F \text{ is a functor} \\ &= \varphi_Y(gf) && \text{by definition of } \varphi \\ &\varphi_Y \circ g_*(f) && \text{by definition of } g_*. \end{aligned}$$

This shows that the diagram above commutes, and we conclude that the assignment φ given by $\varphi_X(f) = F(f)u$ is indeed a natural transformation. We have shown that our proposed map γ is a bijection.

We now have two naturality statements to prove. Naturality in the functor means that given a natural isomorphism $\eta\colon F \longrightarrow G$, the following diagram must commute:

$$\begin{array}{ccc} \mathrm{Nat}(\mathrm{Hom}_{\mathscr{C}}(A,-),F) & \xrightarrow{\gamma_F} & F(A) \\ {\scriptstyle \eta_*}\downarrow & & \downarrow{\scriptstyle \eta_A} \\ \mathrm{Nat}(\mathrm{Hom}_{\mathscr{C}}(A,-),G) & \xrightarrow[\gamma_G]{} & G(A) \end{array}$$

Given a natural transformation φ between $\mathrm{Hom}_{\mathscr{C}}(A,-)$ and F,

$$\begin{aligned} \eta_A\circ\gamma_F(\varphi) &= \eta_A(\varphi_A(1_A)) && \text{by definition of } \gamma \\ &= (\eta\circ\varphi)_A(1_A) && \text{by definition of composition of natural transformations} \\ &= \gamma_G(\eta\circ\varphi) && \text{by definition of } \gamma \\ &= \gamma_G\circ\eta_*(\varphi) && \text{by definition of } \eta_* \end{aligned}$$

so commutativity does hold. Naturality on the object means that given an arrow $f\colon A \longrightarrow B$, the diagram

$$\begin{array}{ccc} \mathrm{Nat}(\mathrm{Hom}_{\mathscr{C}}(A,-),F) & \xrightarrow{\gamma_A} & F(A) \\ {\scriptstyle (f^*)^*}\downarrow & & \downarrow{\scriptstyle F(f)} \\ \mathrm{Nat}(\mathrm{Hom}_{\mathscr{C}}(B,-),F) & \xrightarrow[\gamma_B]{} & F(B) \end{array}$$

commutes. Given a natural transformation φ between $\mathrm{Hom}_{\mathscr{C}}(A,-)$ and F,

$$F(f)\circ\gamma_A(\varphi) = F(f)(\varphi_A(1_A)),$$

while

$$\gamma_B\circ(f^*)^*(\varphi) = \gamma_B(\varphi\circ f^*) = (\varphi\circ f^*)_B(1_B).$$

Now notice that

$$\begin{array}{ccccc} \mathrm{Hom}_{\mathscr{C}}(B,B) & \xrightarrow{f^*} & \mathrm{Hom}_{\mathscr{C}}(A,B) & \xrightarrow{\varphi_B} & F(B) \\ 1_B & \longmapsto & f & \longmapsto & \varphi_B(f) \end{array}.$$

Let's look back at the big commutative diagram we started our proof with: it says in particular that $\varphi_B(f) = F(f)(\varphi_A(1_A))$. So commutativity does hold, and we are done. □

One can naturally (pun intended) define the notion of functor category of contravariant functors, and then prove the corresponding Yoneda Lemma, which will instead use the contravariant Hom functor.

Exercise 1.61 (Contravariant version of the Yoneda Lemma)**.** Let $\mathscr{C}$ be a locally small category, and fix an object B in $\mathscr{C}$. Let $F : \mathscr{C} \longrightarrow \mathbf{Set}$ be a contravariant functor. Then there is a bijection

$$\mathrm{Nat}\ (\mathrm{Hom}_{\mathscr{C}}(-,B), F) \xrightarrow{\gamma} F(B)$$

which is natural on both B and F.

The Yoneda Lemma says that to give a natural transformation between the functors $\mathrm{Hom}_{\mathscr{C}}(A,-)$ and F is choosing an element in the set $F(A)$.

Remark 1.62. Notice that the Yoneda Lemma says in particular that the collection of all natural transformations from $\mathrm{Hom}_{\mathscr{C}}(A,-)$ to F is a set. This wasn't clear a priori, since the collection of objects in $\mathscr{C}$ is not necessarily a set.

The Yoneda Lemma says that natural transformations between representable functors correspond to arrows between the representing objects.

Remark 1.63. If we apply the Yoneda Lemma to the case when F itself is also a Hom functor, say $F = \mathrm{Hom}_{\mathscr{C}}(B,-)$, the Yoneda Lemma says that there is a bijection between $\mathrm{Nat}(\mathrm{Hom}_{\mathscr{C}}(A,-), \mathrm{Hom}_{\mathscr{C}}(B,-))$ and $\mathrm{Hom}_{\mathscr{C}}(B,A)$. In particular, each arrow in $\mathscr{C}$ determines a natural transformation between Hom functors.

The Yoneda Embedding, which we will prove next, formalizes the remark above. It roughly says that every locally small category can be embedded into the category of contravariant functors from $\mathscr{C}$ to **Set**. It is common to refer to both Theorem 1.60 and Theorem 1.65 as the Yoneda Lemma.

Remark 1.64. Let $\mathscr{C}$ be a locally small category. Each arrow $f : A \to B$ in $\mathscr{C}$ gives rise to a natural transformation $\mathrm{Hom}_{\mathscr{C}}(-,A) \implies \mathrm{Hom}_{\mathscr{C}}(-,B)$ that sends each object X to the arrow (function)

$$\begin{array}{ccc} \mathrm{Hom}_{\mathscr{C}}(X,A) & \xrightarrow{f_*} & \mathrm{Hom}_{\mathscr{C}}(X,B) \\ g & \longmapsto & fg. \end{array}$$

The fact that this is a natural transformation is encoded in the following commutative diagram; we have one such diagram for each arrow $g\colon X \to Y$.

$$\begin{array}{ccc} X & & \mathrm{Hom}_{\mathscr{C}}(X,A) \xrightarrow{f_*} \mathrm{Hom}_{\mathscr{C}}(X,B) \\ {\scriptstyle g}\downarrow & & \uparrow{\scriptstyle \mathrm{Hom}_{\mathscr{C}}(g,A)=g^*} \qquad\qquad \uparrow{\scriptstyle \mathrm{Hom}_{\mathscr{C}}(g,B)=g^*} \\ Y & & \mathrm{Hom}_{\mathscr{C}}(Y,A) \xrightarrow[f_*]{} \mathrm{Hom}_{\mathscr{C}}(Y,B) \end{array}$$

This diagram commutes since

$$g^* f_*(h) = g^*(fh) = (fh)g = f(hg) = f_*(hg) = f_* g^*(h).$$

Conversely, f^* indicates the natural transformation $\mathrm{Hom}_{\mathscr{C}}(B,-) \implies \mathrm{Hom}_{\mathscr{C}}(A,-)$ sending each object X to the arrow (function)

$$\begin{array}{c} \mathrm{Hom}_{\mathscr{C}}(B,X) \xrightarrow{f^*} \mathrm{Hom}_{\mathscr{C}}(A,X) \\ g \longmapsto gf. \end{array}$$

Theorem 1.65 (Yoneda Embedding)**.** *Let $\mathscr{C}$ be a locally small category. The covariant functor*

$$\begin{array}{ccc} \mathscr{C} & \longrightarrow & \boldsymbol{Set}^{\mathscr{C}^{op}} \\ A & & \mathrm{Hom}_{\mathscr{C}}(-,A) \\ {\scriptstyle f}\downarrow & \longmapsto & \downarrow{\scriptstyle f_*} \\ B & & \mathrm{Hom}_{\mathscr{C}}(-,B) \end{array}$$

from $\mathscr{C}$ to the category of contravariant functors $\mathscr{C} \longrightarrow \boldsymbol{Set}$ is an embedding. Moreover, the contravariant functor

$$\begin{array}{ccc} \mathscr{C} & \longrightarrow & \boldsymbol{Set}^{\mathscr{C}} \\ A & & \mathrm{Hom}_{\mathscr{C}}(A,-) \\ {\scriptstyle f}\downarrow & \longmapsto & \uparrow{\scriptstyle f^*} \\ B & & \mathrm{Hom}_{\mathscr{C}}(B,-) \end{array}$$

from the category $\mathscr{C}$ to the category of covariant functors $\mathscr{C} \longrightarrow \boldsymbol{Set}$ is also an embedding.

Proof. First, note that our functors are injective on objects because the Hom-sets in our category are all disjoint. So all we need to check is that given objects A and B in $\mathscr{C}$, we have bijections

$$\mathrm{Hom}_{\mathscr{C}}(A,B) \cong \mathrm{Nat}(\mathrm{Hom}_{\mathscr{C}}(-,A), \mathrm{Hom}_{\mathscr{C}}(-,B))$$

and

$$\mathrm{Hom}_{\mathscr{C}^{\mathrm{op}}}(A,B) \cong \mathrm{Nat}(\mathrm{Hom}_{\mathscr{C}}(A,-), \mathrm{Hom}_{\mathscr{C}}(B,-)).$$

Note that the left hand side are the Hom-sets in $\mathscr{C}$, and the right hand side are Hom-sets in **Set**. We will do the details for the second one, and leave the first as an exercise.

This follows from Theorem 1.63, but let's carefully check the details. First, in Theorem 1.64 we have already checked that each arrow is indeed taken to a natural transformation, so we just need to check injectivity and surjectivy at the level of arrows.

The Yoneda Lemma applied here tells us that each natural transformation φ between $\mathrm{Hom}_{\mathscr{C}}(B,-)$ and $F = \mathrm{Hom}_{\mathscr{C}}(A,-)$ corresponds to an element $u \in F(B) = \mathrm{Hom}_{\mathscr{C}}(A,B)$, which we obtain by taking $u := \varphi_B(1_B)$. The Yoneda Lemma says this correspondence is bijective.

Indeed, we can recover φ from u by taking the natural transformation φ that for each object X in $\mathscr{C}$ has $\varphi_X \colon \mathrm{Hom}_{\mathscr{C}}(B,X) \longrightarrow \mathrm{Hom}_{\mathscr{C}}(A,X)$ given by

$$\varphi_X(f) = \mathrm{Hom}_{\mathscr{C}}(f,A)(u) = f_*(u).$$

This shows surjectivity on arrows. Finally, different arrows f give rise to different natural transformations by applying the resulting natural transformation f_* to the identity arrow 1_A, which takes it to f. This shows injectivity on arrows. □

Finally, the Yoneda Embedding says that you can essentially recover an object in a category by knowing the maps from it or into it.

Theorem 1.66. *Let X and Y be objects in a locally small category $\mathscr{C}$. If $\mathrm{Hom}_{\mathscr{C}}(-,X)$ and $\mathrm{Hom}_{\mathscr{C}}(-,Y)$ are naturally isomorphic, or if $\mathrm{Hom}_{\mathscr{C}}(X,-)$ and $\mathrm{Hom}_{\mathscr{C}}(Y,-)$ are naturally isomorphic, then X and Y are isomorphic objects.*

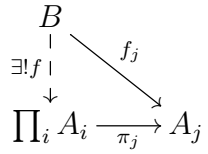
Proof. The Yoneda Embeddings from Theorem 1.65 are fully faithful, and thus by Theorem 1.42 they must reflect isomorphisms. A natural isomorphism between the functors $\mathrm{Hom}_{\mathscr{C}}(X,-)$ and $\mathrm{Hom}_{\mathscr{C}}(Y,-)$ (or the functors $\mathrm{Hom}_{\mathscr{C}}(-,X)$ and $\mathrm{Hom}_{\mathscr{C}}(-,Y)$) is an isomorphism in the target functor category, and it corresponds to f_* (respectively, f^*) for some arrow f from Y to X. By Theorem 1.42, f must be an isomorphism. In particular, X and Y are isomorphic. □

To summarize the content of this chapter, here is the Yoneda Lemma in slogans:

- To give a natural transformation from $\mathrm{Hom}(A,-)$ to F is the same as giving an element in the set $F(A)$.
- The collection of all natural transformations from $\mathrm{Hom}(A,-)$ to F is a set.
- To give a natural transformation between representable functors is to give an arrow between the corresponding representing objects.
- Every locally small category $\mathscr{C}$ can be embedded into the functor category of (covariant or contravariant) functors from $\mathscr{C}$ to **Set**. So rather than studying the category $\mathscr{C}$, we can study functor category to **Set**.
- We can recover an object in a category by knowing the maps from it or into it.

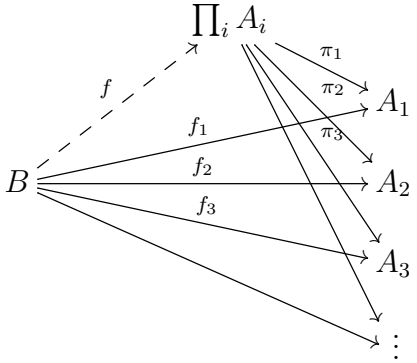
1.5 Products and coproducts

Definition 1.67. Let $\mathscr{C}$ be a locally small category, and consider a family of objects $\{A_i\}_{i\in I}$ in $\mathscr{C}$. The **product** of the A_i is an object in $\mathscr{C}$, denoted by $\prod_i A_i$, together with arrows $\pi_j \in \mathrm{Hom}_{\mathscr{C}}(\prod_i A_i, A_j)$ for each j, called **projections**, satisfying the following universal property: given any object B in $\mathscr{C}$ and arrows $f_i\colon B \longrightarrow A_i$ for each i, there exists a unique arrow f such that

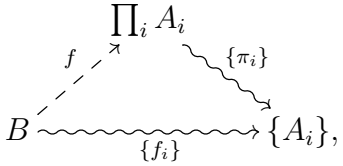


commutes for all j. When I is finite, we may write $A_1 \times \cdots \times A_n$ for the product of $A_1, \ldots, A_n$.

Here is a larger diagram for the (first few) maps involved in a product when the indexing set $I = \mathbb{N}$ is countable:



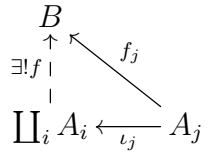
We can also take a "big picture" view of this universal property of the product:



where the squiggly arrows are again collections of maps instead of maps.

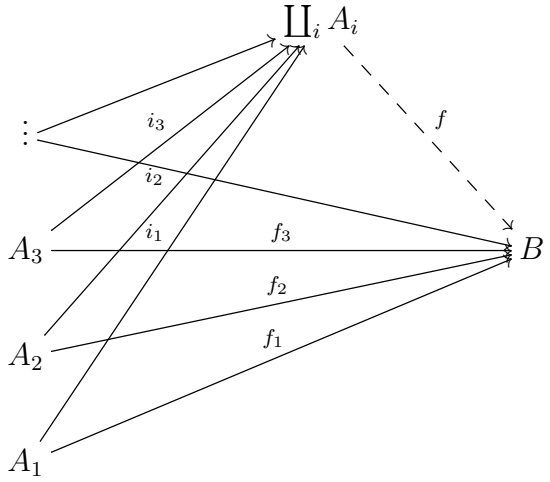
The dual notion is the coproduct.

Definition 1.68. Let $\mathscr{C}$ be a locally small category, and consider a family of objects $\{A_i\}_{i\in I}$ in $\mathscr{C}$. The **coproduct** of the A_i is an object in $\mathscr{C}$, denoted by $\coprod_i A_i$, together with arrows $\iota_j \in \mathrm{Hom}_{\mathscr{C}}(A_j, \coprod_i A_i,)$ for each j, satisfying the following universal property: given any object B in $\mathscr{C}$ and arrows $f_i\colon A_i \longrightarrow B$ for each i, the following diagram commutes:

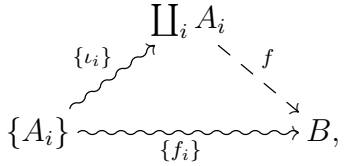


When I is finite, we may write $A_1 \amalg \cdots \amalg A_n$ for the coproduct of $A_1, \ldots, A_n$.

Here is a diagram for the (first few) maps involved in a coproduct when $\Lambda = \mathbb{N}$ is countable:



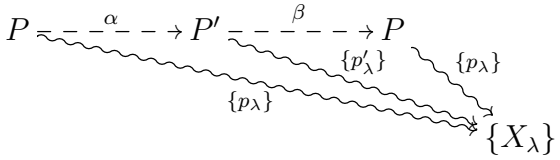
We can also take a "big picture" view of the universal property of the coproduct:



where the squiggly arrows are now collections of maps instead of maps.

Theorem 1.69. *If $(P, \{p_\lambda : P \to X_\lambda\}_{\lambda\in\Lambda})$ and $(P', \{p'_\lambda : P' \to X_\lambda\}_{\lambda\in\Lambda})$ are both products for the same family of objects $\{X_\lambda\}_{\lambda\in\Lambda}$ in a category $\mathscr{C}$, then there is a unique isomorphism $\alpha\colon P \xrightarrow{\sim} P'$ such that $p'_\lambda \circ \alpha = p_\lambda$ for all λ. The analogous statement holds for coproducts.*

Proof. We will just deal with products. The following picture is a rough guide:



Since $(P, \{p_\lambda\})$ is a product and $(P', \{p'_\lambda\})$ is an object with maps to each X_λ, there is a unique map $\beta : P' \to P$ such that $p_\lambda \circ \beta = p'_\lambda$. Switching roles, we obtain a unique map $\alpha : P \to P'$ such that $p'_\lambda \circ \alpha = p_\lambda$.

Consider the composition $\beta \circ \alpha : P \to P$. We have $p_\lambda \circ \beta \circ \alpha = p'_\lambda \circ \alpha = p_\lambda$ for all λ. The identity map $1_P : P \to P$ also satisfies the condition $p_\lambda \circ 1_P = p_\lambda$ for all λ, so by the uniqueness property of products, $\beta \circ \alpha = 1_P$. We can again switch roles to see that $\alpha \circ \beta = 1_{P'}$. Thus α is an isomorphism. The uniqueness of α in the statement is part of the universal property. □

Exercise 1.70. Prove the analogous statement to Theorem 1.69 for coproducts.

This explains why the notations $\prod_i A_i$ and $\coprod_i A_i$ make sense: we can talk about *the* product and *the* coproduct of the A_i, if they exist.

The key thing to remember about these constructions and their universal properties is the following:

- Mapping *into* a product is completely determined by mapping into each of the factors.
- Mapping *out* of a coproduct is completely determined by mapping out of each factor.

Example 1.71. Let $\{X_\lambda\}_{\lambda\in\Lambda}$ be a family of sets. The product of $\{X_\lambda\}_{\lambda\in\Lambda}$ is given by the cartesian product of sets along with the canonical projection maps.

The familiar notion of Cartesian product or direct product serves as a product in many of our favorite categories. Let's note first that given a family of objects $\{X_\lambda\}_{\lambda\in\Lambda}$ in any of the categories $\mathbf{Sgrp}, \mathbf{Grp}, \mathbf{Ring}, R\text{-}\mathbf{Mod}, \mathbf{Top}$, the usual direct product $\prod_{\lambda\in\Lambda} X_\lambda$ is an object of the same category:

- for semigroups, groups, and rings, take the operation coordinate by coordinate:
$$(x_\lambda)_{\lambda\in\Lambda} \cdot (y_\lambda)_{\lambda\in\Lambda} = (x_\lambda \cdot y_\lambda)_{\lambda\in\Lambda};$$
- for modules, addition is coordinate by coordinate, and the action is the same on each coordinate: $r \cdot (x_\lambda)_{\lambda\in\Lambda} = (r \cdot x_\lambda)_{\lambda\in\Lambda}$;
- for topological spaces, use the product topology.

Note that this is not true for fields! The usual product of fields is not a field. In fact, there is no product in this category.

Theorem 1.72. *In each of the categories* **Set**, **Grp**, **Ring**, *R-**Mod**, and* **Top**, *given a family of objects $\{X_\lambda\}_{\lambda\in\Lambda}$, the object $\prod_{\lambda\in\Lambda} X_\lambda$ given by the usual direct product along with the usual projection maps $\pi_\lambda \colon \prod_{\gamma\in\Lambda} X_\gamma \to X_\lambda$ forms a product in the category.*

Proof. We observe that in each category, the direct product is an object, and the projection maps π_λ are morphisms in the category.

Let $\mathscr{C}$ be one of these categories, and suppose that we have morphisms $g_\lambda \colon Y \to X_\lambda$ for all λ in $\mathscr{C}$. We need to show there is a unique morphism $\phi \colon Y \to \prod_{\lambda\in\Lambda} X_\lambda$ such that $\pi_\lambda \circ \phi = g_\lambda$ for all λ. The last condition is equivalent to

$$(\phi(y))_\lambda = (\pi_\lambda \circ \phi)(y) = g_\lambda(y)$$

for all λ, which is equivalent to $\phi(y) = (g_\lambda(y))_{\lambda\in\Lambda}$, so if this is a valid morphism, it is unique. Thus, it suffices to show that the map $\phi(y) = (g_\lambda(y))_{\lambda\in\Lambda}$ is a morphism in $\mathscr{C}$; we leave the details as an exercise. □

Example 1.73. Let $\{X_\lambda\}_{\lambda\in\Lambda}$ be a family of sets. The coproduct of $\{X_\lambda\}_{\lambda\in\Lambda}$ in **Set** is given by the disjoint union with the various inclusion maps. By disjoint union, we simply mean union if the sets are disjoint; in general do something like replace X_λ with $X_\lambda \times \{\lambda\}$ to make them disjoint.

Theorem 1.74. *Let R be a ring, and $\{M_\lambda\}_{\lambda\in\Lambda}$ be a family of left R-modules. A coproduct for the family $\{M_\lambda\}_{\lambda\in\Lambda}$ is given by the direct sum of modules*

$$\bigoplus_{\lambda\in\Lambda} M_\lambda = \{(x_\lambda)_{\lambda\in\Lambda} \mid x_\lambda \neq 0 \text{ for at most finitely many } \lambda\} \subseteq \prod_{\lambda\in\Lambda} M_\lambda$$

together with the inclusion maps

$$M_\lambda \xrightarrow{\iota_\lambda} \bigoplus_{\lambda\in\Lambda} M_\lambda$$

that send each $m \in M_\lambda$ to the tuple that has m in coordinate λ and zeroes elsewhere.

Proof. Given R-module homomorphisms $g_\lambda : M_\lambda \to N$ for each λ, we need to show that there is a unique R-module homomorphism $\alpha : \bigoplus_{\lambda\in\Lambda} M_\lambda \to N$ such that $\alpha \circ \iota_\lambda = g_\lambda$. We define

$$\alpha((m_\lambda)_{\lambda\in\Lambda}) = \sum_{\lambda\in\Lambda} g_\lambda(m_\lambda).$$

Note that since $(m_\lambda)_{\lambda\in\Lambda}$ is in the direct sum, at most finitely many m_λ are nonzero, so the sum on the right hand side is finite, and hence makes sense in N. We need to check that α is R-linear; indeed,

$$\begin{aligned}\alpha((m_\lambda) + (n_\lambda)) &= \alpha((m_\lambda + n_\lambda)) \\ &= \sum g_\lambda(m_\lambda + n_\lambda) \\ &= \sum g_\lambda(m_\lambda) + \sum g_\lambda(n_\lambda) \\ &= \alpha((m_\lambda)) + \alpha((n_\lambda)),\end{aligned}$$

and the check for scalar multiplication is similar. For uniqueness of α, note that $\bigoplus_{\lambda\in\Lambda} M_\lambda$ is generated by the elements $\iota_\lambda(m_\lambda)$ for $m_\lambda \in M_\lambda$. Thus, if α' also satisfies $\alpha' \circ \iota_\lambda = g_\lambda$ for all λ, then $\alpha(\iota_\lambda(m_\lambda)) = g_\lambda(m_\lambda) = \alpha'(\iota_\lambda(m_\lambda))$ so the maps must be equal. □

Remark 1.75. If the index set Λ is finite, then the objects $\prod_{\lambda\in\Lambda} M_\lambda$ and $\bigoplus_{\lambda\in\Lambda} M_\lambda$ are identical, but the product and coproduct are not the same since one involves projection maps and the other involves inclusion maps. When Λ is infinite, the two objects are truly distinct, and in fact the direct sum is a submodule of the product.

Remark 1.76. For any indexing set Λ, $\oplus_{\lambda\in\Lambda} R$ is a free R-module. If $R = k$ happens to be a field, then $\prod_{\lambda\in\Lambda} k$ is free, since all vector spaces are free modules, but in general, $\prod_{\lambda\in\Lambda} R$ is not free for an infinite set Λ.

Example 1.77.

1) In **Top**, disjoint unions serve as coproducts.

2) In **Sgrp** and **Grp**, coproducts exist, and are given as free products. You may see or have seen them in topology in the context of Van Kampen's theorem.

3) In **Ring**, the story is more complicated. Let's note first that disjoint unions won't work, since they are not rings. Direct sums of infinitely many rings do not have 1, so they are not rings in this class, but even finite direct sums or products will not work, since the inclusion maps does not send 1 to 1. We will later on construct coproducts in the full subcategory of **Ring** consisting of commutative rings.

1.6 Limits and colimits

Definition 1.78. Let $(I, \geqslant)$ be a partially ordered set and let $\mathscr{C}$ be a category. An **inverse system** in $\mathscr{C}$ indexed by I is a contravariant functor $\mathbf{PO}(I) \to \mathscr{C}$.

Remark 1.79. Let's unwrap the definition of inverse system a bit. For each $i \in I$, we get an object M_i in $\mathscr{C}$. Moreover, in the category $\mathbf{PO}(I)$, there is exactly one arrow $i \to j$ for each $i \leqslant j$, and the image of this arrow under any contravariant functor $\mathbf{PO}(I) \to \mathscr{C}$ is an arrow $M_j \to M_i$. Finally, our functor must preserve compositions of arrows, so whenever $k \geqslant j \geqslant i$, the arrow $M_k \to M_i$ should match the composition of arrows through j. Thus an inverse system in $\mathscr{C}$ indexed by I consists of the following data:

- for each $i \in I$, an object M_i in $\mathscr{C}$, and
- for each $i \leqslant j$, an arrow $\varphi_i^j \colon M_j \to M_i$ in $\mathscr{C}$

such that whenever $i \leqslant j \leqslant k$, the following diagram must commute:

$$\begin{array}{ccc} M_k & \xrightarrow{\varphi_i^k} & M_i \\ & {\scriptstyle \varphi_j^k} \searrow \quad \nearrow {\scriptstyle \varphi_i^j} & \\ & M_j & \end{array}$$

Note moreover that $\varphi_i^i = \mathrm{id}_{M_i}$, since functors preserve identities. To indicate all this data in a compact way, we say that $\{M_i, \varphi_i^j\}$ is an inverse system.

Example 1.80.

a) An inverse system in a category $\mathscr{C}$ indexed by $\mathbb{N}$ is determined by a diagram of the form

$$X_0 \xleftarrow{a_0} X_1 \xleftarrow{a_1} X_2 \xleftarrow{a_2} X_3 \xleftarrow{a_3} X_4 \xleftarrow{a_4} X_5 \leftarrow \cdots .$$

All the other arrows $X_j \to X_i$ for $i < j$ are given by composition.

b) Let I be a family of submodules of an R-module M. Then we can think of I as a partially ordered set with the reverse inclusion $\supseteq$, so that $L \leqslant N$ if and only if $L \supseteq N$. Whenever $N \subseteq L$, we have an inclusion map $N \to L$, and the family of submodules I together with the inclusion maps forms an inverse system of R-modules.

A special case of this is when we have a descending chain of submodules of M

$$M_1 \supseteq M_2 \supseteq M_3 \supseteq \cdots$$

which is also a special case of an inverse system indexed by $\mathbb{N}$.

c) If I is a poset with the **discrete partial order**, meaning $i \leqslant j$ if and only if $i = j$, then an inverse system indexed by I is just a family of objects indexed by I.

d) If $I = \{1, 2, 3\}$ is a poset with $1 \leqslant 2$ and $1 \leqslant 3$, then an inverse system indexed by I is just a diagram of the form

$$\begin{array}{ccc} & & B \\ & & \downarrow {\scriptstyle f} \\ C & \xrightarrow[g]{} & A. \end{array}$$

Exercise 1.81. Let J be an ideal in a commutative ring R, and consider its nth power, which is the ideal

$$J^n := (f_1 \cdots f_n \mid f_i \in J)$$

generated by all n-fold products of elements in R. For each $m \geqslant n$, consider the maps

$$\begin{aligned} R/J^m &\xrightarrow{\varphi_n^m} R/J^n \\ r + J^m &\longmapsto r + J^n. \end{aligned}$$

Show that these form an inverse system in R-**Mod** indexed by $\mathbb{N}_{>0}$. Note that this can be represented as

$$R/J \xleftarrow{\varphi_1^2} R/J^2 \xleftarrow{\varphi_2^3} R/J^3 \xleftarrow{\varphi_3^4} R/J^4 \xleftarrow{\varphi_4^5} R/J^5 \longleftarrow \cdots.$$

Definition 1.82. Let $\mathscr{C}$ be a category and let $\{M_i, \varphi_i^j\}_i$ be an inverse system on $\mathscr{C}$ indexed by I. The **limit** or **inverse limit** of $\{M_i, \varphi_i^j\}$ consists of an object

$$\varprojlim M_i$$

and arrows

$$\pi_i \colon \varprojlim M_i \to M_i$$

called **projections** such that for all $j \geqslant k$ in I, the diagram

$$\begin{array}{ccc} M_k & \xleftarrow{\pi_k} & \varprojlim M_i \\ {\scriptstyle \varphi_k^j}\uparrow & \swarrow_{\pi_j} & \\ M_j & & \end{array}$$

commutes, and that satisfy the following universal property: for all arrows $f_i \colon X \to M_i$ such that $\varphi_i^j f_j = f_i$ for all i, j, meaning that the diagram

$$\begin{array}{ccc} M_i & \xleftarrow{f_i} & X \\ {\scriptstyle \varphi_i^j}\uparrow & \swarrow_{f_j} & \\ M_j & & \end{array}$$

commutes, there exists a unique arrow $f \colon X \to \varprojlim M_i$ such that

$$\begin{array}{ccccc} \varprojlim M_i & \leftarrow\!\!- & \overset{\exists! f}{-} & -\!\!- & X \\ & {\scriptstyle \pi_j}\searrow & & \swarrow{\scriptstyle f_j} & \\ & & M_j & & \end{array}$$

commute for all j.

One can show that if it exists, the object $\varprojlim M_i$ is unique up to isomorphism; in fact, this is the terminal object in some appropriate (and technical) category. So we can refer to *the* limit of an inverse system. The notation $\varprojlim M_i$ is sometimes replaced by $\lim_i M_i$.

Remark 1.83. Given an inverse system $\{M_i, \varphi_i^j\}$ indexed by I in a category $\mathscr{C}$, say corresponding to the contravariant functor $\varphi : I \to \mathscr{C}$, suppose that its limit exists, and let $L = \varprojlim M_i$. The projections π_i give us commutative diagrams

$$\begin{array}{ccc} L & \xrightarrow{1_L} & L \\ {\scriptstyle \pi_i}\downarrow & & \downarrow{\scriptstyle \pi_j} \\ M_i & \xrightarrow[\varphi_i^j]{} & M_j \end{array}$$

This is the same data as a natural transformation

$$\mathbf{PO}(I) \underset{\varphi}{\overset{\Delta L}{\Downarrow}} \mathscr{C} .$$

In other words, a limit for α consists of an object and a natural transformation from the constant functor on that object to the functor α.

Example 1.84. A terminal object can be viewed as a limit of the empty diagram: since there are no objects in an inverse limit from the empty category, the limit is just an object L that must satisfy the condition that for every object X, there is a unique arrow $X \to L$.

Exercise 1.85. Show that if I is a partially ordered set with the discrete order, then the limit of any inverse system indexed by I is the product on the corresponding set of objects.

Theorem 1.86. *Let R be any ring. Every inverse system of left R-modules over any partially ordered set has a limit.*

Proof. Let I be a partially ordered set and consider an inverse system of R-modules indexed by I, say with modules M_i and homomorphisms $\varphi_i^j \colon M_j \to M_i$. Let

$$L := \{(m_i) \in \prod_i M_i \mid \varphi_i^j(m_j) = m_i \text{ for all } i \leqslant j\}.$$

One can show (exercise!) that this is a submodule of the product of the M_i. For each i, let $\pi_i \colon L \to M_i$ be the restriction of the projection maps $\prod M_i \to M_i$ to L. We claim that L is a limit for the inverse system, together with the projection maps π_i.

First, note that

$$\varphi_i^j \pi_j((m_k)_k) = \varphi_i^j(m_j) = m_i = \pi_i((m_k)_k),$$

by construction, so $\varphi_i^j \pi_j = \pi_i$.

Moreover, suppose that we are given an R-module X and R-module homomorphisms $f_i \colon X \to M_i$ such that $\varphi_i^j f_j = f_i$ for all $i \leqslant j$. Define

$$\begin{array}{ccc} X & \xrightarrow{g} & \prod_i M_i \\ x & \longmapsto & (f_i(x))_i. \end{array}$$

First, note that $\pi_i(g(x)) = f_i(x)$ for all i by construction. Moreover, this is an R-module homomorphism; it is induced by the universal property of the product. We claim that the

image of g is contained in L, and thus that we can restrict g to an R-module homomorphism $f\colon X \to L$. Indeed, given any $x \in X$,

$$\varphi_i^j(\pi_j(g(x))) = \varphi_i^j(f_j(x)) = f_i(x) = \pi_i(g(x)).$$

This says that $g(x) \in L$, so we get an R-module homomorphism $f\colon X \to L$ given by

$$f(x) = (f_i(x))_i.$$

Finally, we claim that L and f satisfy the desired universal property, and for that, we need first to check that

$$\begin{array}{ccccc} \varprojlim M_i & \xleftarrow{\quad} & \cdots & \xleftarrow{f} & X \\ & \searrow^{\pi_i} & & \swarrow_{f_i} & \\ & & M_i & & \end{array}$$

commutes, and we need to check that such f is unique. The commutativity is immediate, since as noted above $\pi(f(x)) = f_i(x)$ for all $x \in X$ by construction. For uniqueness, suppose that h is any other R-module homomorphism $X \to L$ such that

$$\begin{array}{ccccc} \varprojlim M_i & \xleftarrow{\quad} & \cdots & \xleftarrow{h} & X \\ & \searrow^{\pi_i} & & \swarrow_{f_i} & \\ & & M_i & & \end{array}$$

also commutes. Given any $x \in X$, let $h(x) = (m_i)$. Then

$$m_i = \pi_i(h(x)) = f_i(x)$$

for all i, so

$$h(x) = (m_i)_i = (f_i(x))_i = f(x),$$

and thus $h = f$. This completes the proof that L is a limit for the given inverse system. □

Remark 1.87. One can adapt the proof of Theorem 1.86 to show that all limits in **Set** exist, and can be constructed explicitly as a subset of the product of the sets forming the inverse system: the limit of an inverse system $\{M_i, \varphi_i^j\}$ is the subset of the product given by

$$L := \{(m_i) \in \prod_i M_i \mid \varphi_i^j(m_j) = m_i \text{ for all } i \leqslant j\}$$

together with the canonical projections from the product restricted to the subset L.

Example 1.88.

a) If I is a partially ordered set with the discrete order, then the limit of any inverse system just the product.

b) Given a ring R and an ideal J, the limit of the inverse system

$$R/J \longleftarrow R/J^2 \longleftarrow R/J^3 \longleftarrow R/J^4 \longleftarrow R/J^5 \longleftarrow \cdots$$

is the J-adic completion of R.

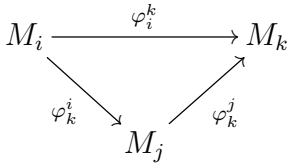
The dual construction to limits is the notion of a colimit.

Definition 1.89. Let $(I, \geqslant)$ be a partially ordered set and let $\mathscr{C}$ be a category. A **direct system** in $\mathscr{C}$ indexed by I is a covariant functor $\mathbf{PO}(I) \to \mathscr{C}$.

Remark 1.90. An inverse system in $\mathscr{C}$ indexed by I consists of the following data:

- for each $i \in I$, an object M_i in $\mathscr{C}$, and
- for each $i \leqslant j$, an arrow $\varphi^i_j \colon M_i \to M_j$ in $\mathscr{C}$

such that whenever $i \leqslant j \leqslant k$, the following diagram must commute:



Note moreover that $\varphi^i_i = \mathrm{id}_{M_i}$, since functors preserve identities. To indicate all this data in a compact way, we say that $\{M_i, \varphi^j_i\}$ is an inverse system.

Example 1.91.

a) A direct system in a category $\mathscr{C}$ indexed by $\mathbb{N}$ is determined by a diagram of the form

$$X_1 \xrightarrow{a_1} X_2 \xrightarrow{a_2} X_3 \xrightarrow{a_3} X_4 \xrightarrow{a_4} X_5 \to \cdots .$$

All the other arrows $X_i \to X_j$ for $i < j$ are given by composition.

b) Let I be a family of submodules of an R-module M. Then we can think of I as a partially ordered set with $\subseteq$. Whenever $N \subseteq L$, we have an inclusion map $N \to L$, and the family of submodules I together with the inclusion maps forms a direct system of R-modules.

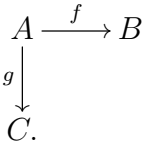
A special case of this is when we have an ascending chain of submodules of M

$$M_1 \subseteq M_2 \subseteq M_3 \subseteq \cdots$$

which is also a special case of a direct system indexed by $\mathbb{N}$.

c) If I is a poset with the discrete partial order, then an inverse system indexed by I is just a family of objects indexed by I.

d) If $I = \{1, 2, 3\}$ is a poset with $1 \leqslant 2$ and $1 \leqslant 3$, then a direct system indexed by I is just a diagram of the form



Definition 1.92. Let $\mathscr{C}$ be a category and let $\{M_i, \varphi_j^i\}_i$ be a direct system on $\mathscr{C}$ indexed by I. The **colimit** or **direct limit** of $\{M_i, \varphi_j^i\}$ consists of an object

$$\varinjlim M_i$$

and arrows

$$\alpha_i \colon M_i \to \varinjlim M_i$$

called **insertion arrows** such that

$$\alpha_j \varphi_j^i = \alpha_i \quad \text{for all } i, j \in I$$

satisfying the following universal property: for all arrows $f_i \colon M_i \to X$ such that $f_j \varphi_j^i = f_i$ for all i, j, meaning that the diagram

$$\begin{array}{ccc} M_i & \xrightarrow{f_i} & X \\ {\scriptstyle \varphi_j^i}\downarrow & \nearrow_{f_j} & \\ M_j & & \end{array}$$

commutes, there exists a unique arrow $f \colon \varinjlim M_i \to X$ such that

$$\begin{array}{ccccc} \varinjlim M_i & & \overset{\exists! f}{- - - \to} & & X \\ & {\scriptstyle \alpha_j}\nwarrow & & \nearrow{\scriptstyle f_j} & \\ & & M_j & & \end{array}$$

commutes.

One can show that if it exists, the object $\varinjlim M_i$ is unique up to isomorphism; in fact, this is the initial object in some appropriate (and technical) category. So we can refer to *the* colimit of a direct system. The notation $\varinjlim M_i$ is sometimes replaced by $\operatorname{colim}_i M_i$.

Remark 1.93. Given a direct system $\{M_i, \varphi_j^i\}$ indexed by I in a category $\mathscr{C}$, say corresponding to the covariant functor $\varphi \colon I \to \mathscr{C}$, suppose that its colimit exists, and let $L = \varinjlim M_i$. The α_i give us commutative diagrams

$$\begin{array}{ccc} L & \xrightarrow{1_L} & L \\ {\scriptstyle \alpha_i}\uparrow & & \uparrow{\scriptstyle \alpha_j} \\ M_i & \xrightarrow[\varphi_j^i]{} & M_j \end{array}$$

This is the same data as a natural transformation

$$\mathscr{C} \underset{\Delta L}{\overset{\varphi}{\Downarrow}} \mathscr{C} \,.$$

In other words, a limit for α consists of an object and a natural transformation from α to the constant functor on that object.

Example 1.94. An initial object can be viewed as a colimit of the empty diagram: since there are no objects in a direct limit from the empty category, the colimit is an object C that must satisfy the condition that for every object X, there is a unique arrow $C \to X$.

Exercise 1.95. Show that if I is a poset with the discrete order, then the colimit of any inverse system indexed by I is the same as the coproduct of the corresponding set of objects.

Theorem 1.96. *Let R be any ring. Every direct system of left R-modules over any partially ordered has a colimit.*

Proof. Let I be a partially ordered set and consider a direct system of R-modules indexed by I, say with modules M_i and homomorphisms $\varphi^i_j \colon M_j \to M_i$. Let $\iota_i \colon M_i \to \oplus_j M_j$ be the inclusions into the direct sum, let S be the submodule of $\oplus M_i$ generated by all elements of the form

$$\iota_i(\varphi^i_j(m_i)) - \iota_i(m_i),$$

and define

$$C := \bigoplus_i M_i/S.$$

For each i, let

$$\begin{aligned} M_i &\xrightarrow{\quad\alpha_i\quad} C \\ m &\longmapsto \iota_i(m) + S. \end{aligned}$$

We claim that C together with the maps α_i is a colimit for the direct system; we leave the details as an exercise. □

Remark 1.97. One can adapt the proof of Theorem 1.86 to show that all colimits in **Set** exist, and can be constructed explicitly as the set of equivalence classes of an appropriate equivalence relation on the coproduct.

There are many other important constructions that arise as special cases of limits and colimits, some of which we will study later in the class. Here is one more example:

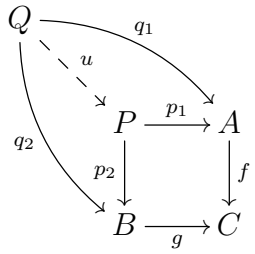
Definition 1.98. Let $\mathscr{C}$ be a category. A **pullback** of the arrows f and g consists of an object P and arrows p_1 and p_2 such that

$$\begin{array}{ccc} P & \xrightarrow{p_1} & A \\ {\scriptstyle p_2}\downarrow & & \downarrow{\scriptstyle f} \\ B & \xrightarrow[g]{} & C \end{array}$$

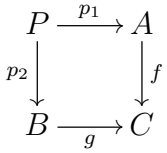
commutes, and satisfying the following universal property: for all objects Q and arrows q_1 and q_2 such that

$$\begin{array}{ccc} Q & \xrightarrow{q_1} & A \\ {\scriptstyle q_2}\downarrow & & \downarrow{\scriptstyle f} \\ B & \xrightarrow[g]{} & C \end{array}$$

commutes, there exists a unique u such that

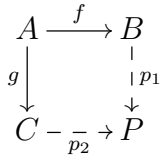


commutes. One sometimes refers to the following diagram as a **pullback diagram**:

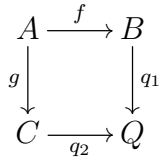


The dual construction is the pushout.

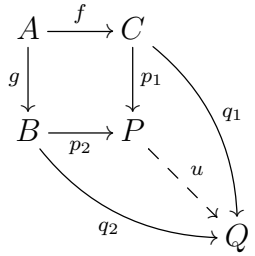
Definition 1.99. Let $\mathscr{C}$ be a category. A **pushout** of the arrows f and g consists of an object P and arrows p_1 and p_2 such that



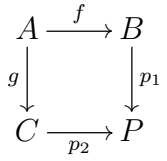
commutes, and satisfying the following universal property: for all objects Q and arrows q_1 and q_2 such that



commutes, there exists a unique u such that



commutes. One sometimes refers to the following diagram as a **pushout diagram**:



Exercise 1.100. Interpret the notion of pullback as a limit and a pushout as a colimit. More precisely, describe a partially ordered set and corresponding inverse system or direct system whose limit or colimit is the same as a pushout or pullback.

We showed in Theorem 1.86 and Theorem 1.96 that R-**Mod** has all limits and colimits. In the case of pullbacks and pushouts, one can describe the corresponding module in a more manageable way.

Exercise 1.101. Explicitly describe pullbacks and pushouts in R-**Mod**.

We have defined the limit of an inverse system, and the colimit of a direct system. One can define limits and colimits even more generally – every functor may have a limit, even if the source is not a poset category.

Definition 1.102. Let $\mathscr{C}$ be a category and let J be a small category. A **diagram** in $\mathscr{C}$ of **shape** J is a functor $J \to \mathscr{C}$. We may call J the **index category**.

Remark 1.103. Let $\mathscr{C}$ be a category and let J be a small category. Let I be the set of objects in J. To give a covariant functor $J \to \mathscr{C}$ is to give

- a set $\{X_i\}_{i \in I}$ in $\mathscr{C}$ indexed by I,
- or every pair (i, j) of objects $i, j \in I$, a set of arrows $A_{i,j} := \{f_\alpha\}$ in $\mathrm{Hom}_{\mathscr{C}}(X_i, X_j)$ indexed by the set $\mathrm{Hom}_I(i, j)$

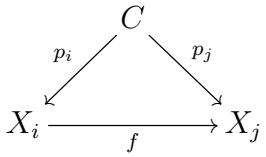
satisfying the necessary properties to guarantee that that 1_i gets sent to 1_{X_i} and that composition of arrows is preserved. One can give a diagram by forgetting the underlying indexing category J and just presenting the set of objects, sets of arrows, and corresponding composition rules.

One advantage of giving this data, as opposed to the functor $F : J \to \mathscr{C}$, is that we do not need to distinguish between covariant and contravariant functors – we are simply giving a set of objects and various sets of arrows.

Definition 1.104. Consider a diagram in $\mathscr{C}$ with objects $\{X_i\}_{i \in I}$ in $\mathscr{C}$ and arrows $A_{i,j} = \{f_\alpha\}$. A **cone** over this diagram consists of

- an object C in $\mathscr{C}$, and
- for each $i \in I$, an arrow $p_i : C \to X_i$

such that for every pair (i, j) and every arrow $f : X_i \to X_j$ in the diagram, the following triangle commutes:

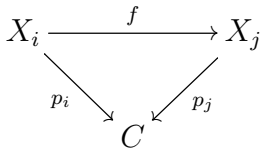


Dually, a **cocone** over this diagram consists of

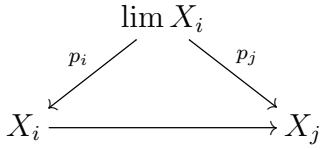
- an object C in $\mathscr{C}$, and

- for each $i \in I$, an arrow $p_i \colon X_i \to C$

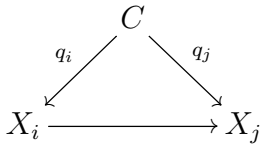
such that for every pair (i, j) and every arrow $f : X_i \to X_j$ in the diagram, the following triangle commutes:



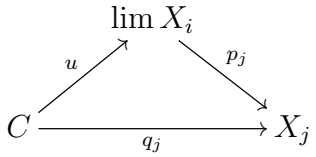
Definition 1.105. Consider a diagram in $\mathscr{C}$ with objects $\{X_i\}_{i \in I}$ in $\mathscr{C}$ and arrows $A_{i,j} = \{f_\alpha\}$. The **limit** of this diagram is, if it exists, a cone



which is terminal with respect to all other cones, meaning that for every other cone

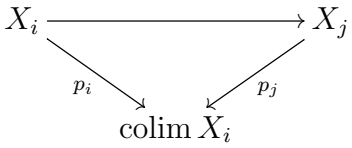


there exists a unique arrow $u \colon C \to \lim X_i$ such that

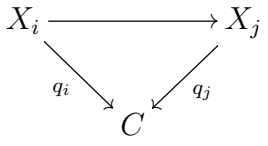


commutes.

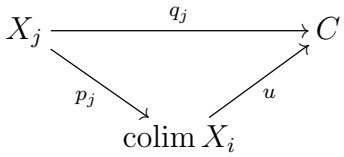
The **colimit** of this diagram is, if it exists, a cocone



which is initial with respect to all other cones, meaning that for every other cone



there exists a unique arrow $u \colon \operatorname{colim} X_i \to C$ such that



commutes.

One can check that if we take a limit of a contravariant diagram indexed by a poset category, we recover the limit of an inverse system, and analogously the colimit of a covariant diagram indexed by a poset category is the colimit of a direct system.

Definition 1.106. A covariant functor $F\colon \mathscr{C} \to \mathscr{D}$

- **preserves colimits** if
$$F(\operatorname{colim} M_i) = \operatorname{colim} F(M_i).$$
More precisely, if the object $\operatorname{colim} M_i$ and the arrows $\alpha_i : M_i \to \operatorname{colim} M_i$ form the colimit of diagram D, then $F(\operatorname{colim} M_i)$ is the colimit of the diagram $F \circ D$ with insertion arrows $F(\alpha_j)\colon F(M_j) \to F(\operatorname{colim} M_i)$.

- **preserves limits** if
$$F(\lim M_i) = \lim F(M_i).$$
More precisely, if $\lim M_i$ is the limit of a diagram D with projections $\pi_j\colon \lim M_i \to M_j$, then the object $F(\lim M_i)$ and the projection arrows $F(\pi_j)\colon F(\lim M_i) \to F(M_j)$ form a limit of the diagram $F \circ D$.

Definition 1.107. A contravariant functor $F : \mathscr{C} \to \mathscr{D}$ **converts limits to colimits** or **sends limits to colimits** if

$$F(\lim M_i) = \operatorname{colim} F(M_i).$$

Similarly, a contravariant functor $F : \mathscr{C} \to \mathscr{D}$ **converts colimits to limits** or **sends colimits to limits** if
$$F(\operatorname{colim} M_i) = \lim F(M_i).$$

The Hom functors preserve limits and colimits.

Theorem 1.108. *Let $\mathscr{C}$ be any category and let A be an object in $\mathscr{C}$.*

a) If the limit $\lim_i M_i$ exists, then there is a natural isomorphism

$$\operatorname{Hom}_{\mathscr{C}}(A, \lim_i M_i) \cong \lim_i \operatorname{Hom}_{\mathscr{C}}(A, M_i).$$

In particular, the limit of $\operatorname{Hom}_{\mathscr{C}}(A, M_i)$ exists.

b) If the limit $\lim_i M_i$ exists, then there is a natural isomorphism

$$\operatorname{Hom}_{\mathscr{C}}(\operatorname{colim}_i M_i, A) \cong \lim_i \operatorname{Hom}_{\mathscr{C}}(M_i, B).$$

In particular, the limit of $\operatorname{Hom}_{\mathscr{C}}(M_i, A)$ exists.

1.7 Universal properties

We have all seen constructions that are at first a bit messy but that end up satisfying some nice universal property that makes everything work out. At the end of the day, a universal property allows us to ignore the messy details and focus on the universal property, which usually says everything we need to know about the construction.

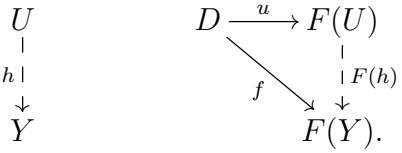
Universal properties are *everywhere*. Limits and colimits are a big example; products and coproducts are a special case of limits and colimits. A representable functor encodes a *universal property* of the object that represents it: for example, in Theorem 1.57, mapping out of the singleton set is the same as choosing an element x in a set X.

In this section, we will briefly describe how one can formalize the idea of a universal property in categorical language. This is not necessary to understand what comes afterwards; this section is here for our own amusement. The most interesting observation in this section is perhaps that any universal property can be phrased in terms of representable functors. There are a few different equivalent frameworks in the literature, and we will briefly try to reconcile two of them. We note, however, that understanding this formalism is not necessarily for what we will do next; this level of abstraction can be confusing at first, and this is a section that can be better understood once the reader has had some time to get comfortable with categorical language.

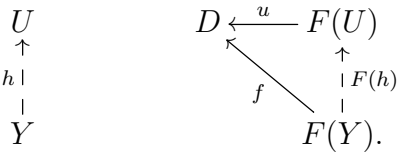
Definition 1.109. Let $\mathscr{C}$ be a locally small category. A **universal property** of an object C in $\mathscr{C}$ consists of a representable functor $F\colon \mathscr{C} \longrightarrow \mathbf{Set}$ together with a **universal element** $X \in F(C)$ such that F is naturally isomorphic to either $\mathrm{Hom}_{\mathscr{C}}(C, -)$ (if F is covariant) or $\mathrm{Hom}_{\mathscr{C}}(-, C)$ (if F is contravariant), via the natural isomorphism that corresponds to X via the bijection in the Yoneda Lemma.

We can rephrase this in terms of universal arrows.

Definition 1.110. Let $F\colon \mathscr{C} \longrightarrow \mathscr{D}$ be covariant functor and let D be an object in $\mathscr{D}$. A **universal arrow from** D **to** F is a pair (U, u) where U is an object in $\mathscr{C}$ and an arrow $u \in \mathrm{Hom}_{\mathscr{D}}(D, F(U))$ with the following **universal property**: for any arrow $f \in \mathrm{Hom}_{\mathscr{D}}(D, F(Y))$, there exists a unique arrow $h \in \mathrm{Hom}_{\mathscr{C}}(U, Y)$ such that the following diagram commutes:



There is a dual to this definition. A **universal arrow from** F **to** D is a pair (U, u), where C is an object in $\mathscr{C}$ and $u \in \mathrm{Hom}_{\mathscr{D}}(F(U), D)$ that satisfy the following **universal property**: for any arrow $f \in \mathrm{Hom}_{\mathscr{D}}(F(Y), D)$, there exists a unique $h \in \mathrm{Hom}_{\mathscr{C}}(Y, U)$ such that the following diagram commutes:



Let's see in detail why it is that giving a universal arrow is equivalent to giving a universal property as defined above.

Remark 1.111. Let $F : \mathscr{C} \longrightarrow \mathscr{D}$ be a covariant functor, and fix an object U in $\mathscr{C}$, an object D in $\mathscr{D}$, and an arrow $u \in \mathrm{Hom}_{\mathscr{D}}(D, F(U))$. Notice that $\mathrm{Hom}_{\mathscr{D}}(D, F(-))$ determines a covariant functor $\mathscr{C} \longrightarrow \mathbf{Set}$. By the Yoneda Lemma, the following is a recipe for a natural transformation between $\mathrm{Hom}_{\mathscr{C}}(U, -)$ and $\mathrm{Hom}_{\mathscr{D}}(D, F(-))$: for each object Y in $\mathscr{C}$ and each arrow $h \in \mathrm{Hom}_{\mathscr{C}}(U, Y)$, set

$$\varphi_Y(h) := \mathrm{Hom}_{\mathscr{D}}(D, F(h))(u).$$

Notice that

$$\begin{array}{ccc} \mathrm{Hom}_{\mathscr{D}}(D, F(U)) & \xrightarrow{\mathrm{Hom}_{\mathscr{D}}(D, F(h))} & \mathrm{Hom}_{\mathscr{D}}(D, F(Y)) \ , \\ f & \longmapsto & F(h) \circ u \end{array}$$

so $\varphi_Y(h)(f) = F(h) \circ u$.

We get the following commutative diagram:

$$\begin{array}{ccc} \mathrm{Hom}_{\mathscr{C}}(U, U) & \xrightarrow{\mathrm{Hom}_{\mathscr{C}}(U, h)} & \mathrm{Hom}_{\mathscr{C}}(U, Y) \\ \Big\downarrow{\scriptstyle \varphi_U} & & \Big\downarrow{\scriptstyle \varphi_Y} \\ 1_U & \longmapsto & h \\ \Big\downarrow & & \Big\downarrow \\ u & \longmapsto & F(h) \circ u =: \varphi_Y(h) \\ \Big\downarrow & & \Big\downarrow \\ \mathrm{Hom}_{\mathscr{D}}(D, F(U)) & \xrightarrow[\mathrm{Hom}_{\mathscr{D}}(D, F(h))]{} & \mathrm{Hom}_{\mathscr{D}}(D, F(Y)) \end{array}$$

Given an arrow $f \in \mathrm{Hom}_{\mathscr{D}}(D, F(Y))$, $\varphi_Y(h) = f$ for some $h \in \mathrm{Hom}_{\mathscr{C}}(U, Y)$ if and only if $F(h) \circ u = f$.

On the one hand, φ is a natural isomorphism if and only if for every object Y in $\mathscr{C}$ and every $f \in \mathrm{Hom}_{\mathscr{D}}(D, F(Y))$ there exists a unique $h \in \mathrm{Hom}_{\mathscr{C}}(U, Y)$ such that $F(h) \circ u = f$. On the other hand, that is exactly the condition required for (U, u) to be a universal arrow from D to F. So we have shown that the following are equivalent:

- (U, u) is a universal arrow from D to F.
- U represents the functor $\mathrm{Hom}_{\mathscr{D}}(D, F(-)) \colon \mathscr{C} \longrightarrow \mathbf{Set}$, via $u \in \mathrm{Hom}_{\mathscr{D}}(D, F(U))$.

Similarly, one can prove the dual equivalence:

- (U, u) is a universal arrow from F to D.
- U represents the functor $\mathrm{Hom}_{\mathscr{D}}(F(-), D) \colon \mathscr{C} \longrightarrow \mathbf{Set}$, via $u \in \mathrm{Hom}_{\mathscr{D}}(F(U), D)$.

Conversely, suppose that we are given a representable functor $F : \mathscr{C} \longrightarrow \mathbf{Set}$ together with an element $X \in F(C)$ such that F is naturally isomorphic to $\mathrm{Hom}_{\mathscr{C}}(C, -)$ via the natural isomorphism that corresponds to X via the bijection in the Yoneda Lemma. First, let $\{\star\}$ be a singleton. Recall that we saw in Theorem 1.57 that the functor $\mathrm{Hom}_{\mathbf{Set}}(\{\star\}, -)$ is naturally isomorphic to the identity functor on **Set**; by composing natural isomorphisms, this implies that $\mathrm{Hom}_{\mathbf{Set}}(\{\star\}, F(-)) = \mathrm{Hom}_{\mathbf{Set}}(\{\star\}, -) \circ F$ is naturally isomorphic to $\mathrm{Hom}_{\mathscr{C}}(C, -)$. So the object C represents the functor $\mathrm{Hom}_{\mathbf{Set}}(\{\star\}, F(-))$; this is half the recipe for a universal arrow.

Now if we actually want to keep track of the arrow $u \in \mathrm{Hom}_{\mathbf{Set}}(\{\star\}, F(C))$ that corresponds to this natural isomorphism, we need to keep track of what happens when we compose with $\mathrm{Hom}_{\mathbf{Set}}(\{\star\}, -)$. We started with a natural isomorphism corresponding to $X \in F(C)$, and composed with the functor $\mathrm{Hom}_{\mathbf{Set}}(\{\star\}, -)$, so our original $X \in F(C)$ will now correspond to some element in $\mathrm{Hom}_{\mathbf{Set}}(\{\star\}, F(C))$; this set is in natural bijection with the original set $F(C)$, and the element $X \in F(C)$ corresponds to the function $u \in \mathrm{Hom}_{\mathbf{Set}}(\{\star\}, F(C))$ given by $\star \mapsto X$. This is the arrow u we are searching for.

In conclusion: we have an equivalence between the following pieces of data:

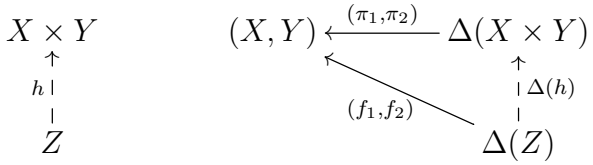
- A representable functor $F : \mathscr{C} \longrightarrow \mathbf{Set}$ together with an element $X \in F(C)$ such that F is naturally isomorphic to $\mathrm{Hom}_{\mathscr{C}}(C, -)$ via the natural isomorphism that corresponds to X via the bijection in the Yoneda Lemma.
- A universal arrow (C, u) from $\{\star\}$ to F, where $u \in \mathrm{Hom}_{\mathbf{Set}}(\{\star\}, F(C))$ is given by $\star \mapsto X$.

Let's take some of the universal properties we have encountered before and try to rephrase them via this formal lens.

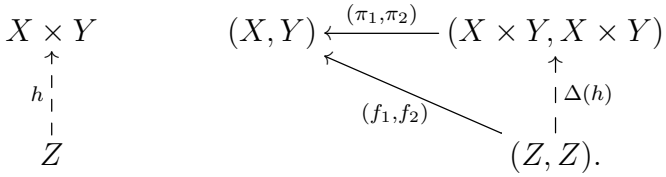
Example 1.112. A singleton set $\{\star\}$ (or *the* singleton set, if we think about sets up to isomorphism) has the following simple universal property: to give a function out of $\{\star\}$ is the same as choosing an element in the target set. We saw in Theorem 1.57 that this is encoded in the fact that the identity functor on **Set** is representable, with representing object $\{\star\}$. Now here is a fun fact: the natural isomorphism between the identity on **Set** and the functor $\mathrm{Hom}_{\mathbf{Set}}(\{\star\}, -)$ used is the *only* natural transformation between them: indeed, the Yoneda Lemma says that each natural transformation corresponds to an element in $1_{\mathbf{Set}}(\{\star\}) = \{\star\}$; but there is only one such element!

Example 1.113. Let's phrase the universal property of products as a universal property in this formal sense, at least in the case of the product of two object C_1 and C_2 in $\mathscr{C}$. To do that, we need to consider the **product category** $\mathscr{C} \times \mathscr{C}$ whose objects are pairs (C_1, C_2) of objects in $\mathscr{C}$, and an arrow $(C_1, C_2) \longrightarrow (C_3, C_4)$ is given by a pair (f_1, f_2) with $f_1 \in \mathrm{Hom}_{\mathscr{C}}(C_1, C_3)$ and $f_2 \in \mathrm{Hom}_{\mathscr{C}}(C_2, C_4)$. The diagonal functor $\Delta : \mathscr{C} \longrightarrow \mathscr{C} \times \mathscr{C}$ is exactly what it sounds like: $\Delta(C) = (C, C)$ for every object C in $\mathscr{C}$ and $\Delta(f) = (f, f)$ for every arrow f in $\mathscr{C}$.

Given objects X and Y in $\mathscr{C}$, consider the projection arrows $\pi_1 : X \times Y \longrightarrow X$ and $\pi_2 : X \times Y \longrightarrow Y$. We claim that the object $X \times Y$ together with the arrow (π_1, π_2) in $\mathscr{C} \times \mathscr{C}$ form a universal arrow from Δ to (X, Y) in $\mathscr{C} \times \mathscr{C}$. Why? If true, this would mean that given any object Z in $\mathscr{C}$ and any arrow $(f_1, f_2) \in \mathrm{Hom}_{\mathscr{C} \times \mathscr{C}}(\Delta(Z), (X, Y))$, there exists a unique $h \in \mathrm{Hom}_{\mathscr{C}}(Z, X \times Y)$ such that



commutes. This is indeed the universal property of products we described less formally when we first defined products: given $f_1 : Z \longrightarrow X$ and $f_2 : Z \longrightarrow Y$, there is a unique $h : Y \longrightarrow X \times Y$ such that the following diagram commutes:



The diagram in $\mathscr{C} \times \mathscr{C}$ translates into two commutative diagrams in $\mathscr{C}$:

This is precisely the universal property of the product that we described before.

Equivalently, following the recipe we described in Theorem 1.111, the universal property of the product is encoded in the representable functor $\mathrm{Hom}_{\mathscr{C}\times\mathscr{C}}(\Delta(-), (X, Y))$, which is represented by $X \times Y$ via (π_1, π_2). So there is a natural isomorphism

$$\mathrm{Hom}_{\mathscr{C}}(-, X \times Y) \cong \mathrm{Hom}_{\mathscr{C}\times\mathscr{C}}(\Delta(-), (X, Y)),$$

which means that to give an arrow to $X \times Y$ is the same as giving an arrow to X and an arrow to Y. In fact, this natural iso is the natural transformation that the Yoneda bijection we constructed in Theorem 1.60 takes to $(\pi_1, \pi_2) \in \mathrm{Hom}_{\mathscr{C}}(\Delta(X \times Y), (X, Y))$. If we follow that bijection, our natural isomorphism φ sends an object Z in $\mathscr{C}$ to the arrow

$$\begin{aligned} \mathrm{Hom}_{\mathscr{C}}(Z, X \times Y) &\xrightarrow{\quad\varphi_Z\quad} \mathrm{Hom}_{\mathscr{C}\times\mathscr{C}}(\Delta(Z), (X, Y)) \\ f &\longmapsto \left(\Delta(Z) \xrightarrow{(f,f)} \Delta(X \times Y) \xrightarrow{(\pi_1,\pi_2)} (X, Y)\right). \end{aligned}$$

Since φ_Z is a bijection, every arrow $(f_1, f_2) \in \mathrm{Hom}_{\mathscr{C}\times\mathscr{C}}(\Delta(Z), (X, Y))$ is $\varphi_Z(f)$ for some $(f_1, f_2) \in \mathrm{Hom}_{\mathscr{C}}(Z, X \times Y)$. In particular, there exists (f_1, f_2) such that $f_1 = \pi_1 f$ and $f_2 = \pi_2 f$. And surprise surprise: we just rediscovered the universal property of the product!

Exercise 1.114. Rephrase the universal property of the coproduct in this formal sense.

1.8 Adjoint functors

Universal properties are closely related to adjoint functors.

Definition 1.115. Let $\mathscr{C}$ and $\mathscr{D}$ be locally small categories. Two covariant functors

$$\mathscr{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathscr{D}$$

form an **adjoint** pair (F, G) if given any objects $C \in \mathscr{C}$ and $D \in \mathscr{D}$, there is a bijection between the Hom-sets

$$\mathrm{Hom}_{\mathscr{D}}(F(C), D) \xrightarrow{\cong} \mathrm{Hom}_{\mathscr{C}}(C, G(D))$$

which is natural on both objects, meaning that for all $f \in \mathrm{Hom}_{\mathscr{C}}(C_1, C_2)$ and $g \in \mathrm{Hom}_{\mathscr{D}}(D_1, D_2)$, the diagrams

$$\begin{array}{ccc} \mathrm{Hom}_{\mathscr{D}}(F(C_1), D) & \xrightarrow{\cong} & \mathrm{Hom}_{\mathscr{C}}(C_1, G(D)) \\ {\scriptstyle F(f)^*}\downarrow & & \downarrow{\scriptstyle f^*} \\ \mathrm{Hom}_{\mathscr{D}}(F(C_2), D) & \xrightarrow{\cong} & \mathrm{Hom}_{\mathscr{C}}(C_2, G(D)) \end{array} \quad \text{and} \quad \begin{array}{ccc} \mathrm{Hom}_{\mathscr{D}}(F(C), D_1) & \xrightarrow{\cong} & \mathrm{Hom}_{\mathscr{C}}(C, G(D_1)) \\ {\scriptstyle g_*}\downarrow & & \downarrow{\scriptstyle G(g)_*} \\ \mathrm{Hom}_{\mathscr{D}}(F(C), D_2) & \xrightarrow{\cong} & \mathrm{Hom}_{\mathscr{C}}(C, G(D_2)) \end{array}$$

commute for all $C \in \mathscr{C}$ and all $D \in \mathscr{D}$. We say that F is the **left adjoint** of G, or that F has a **right adjoint**, and that G is the **right adjoint** of F, or that G has a **left adjoint**.

We can think of adjoint functors as solutions to optimization problems. A particular adjoint functor gives the most efficient functorial solution to some problem.

Example 1.116. Fix a ring R. Given a set I, what is the most efficient way to assign an R-module to I in a functorial way? The solution to this problem is the construction of free modules. Formally, the free functor is the functor $\mathbf{Free}\colon \mathbf{Set} \longrightarrow R\text{-}\mathbf{Mod}$ that sends each set I to the free R-module on I

$$R^I = \bigoplus_I R.$$

The free functor is precisely a left adjoint to the forgetful functor $R\text{-}\mathbf{Mod} \longrightarrow \mathbf{Set}$. That is, there is a natural bijection

$$\mathrm{Hom}_{R\text{-}\mathbf{Mod}}(\bigoplus_I R, M) \cong \mathrm{Hom}_{\mathbf{Set}}(I, M).$$

(On the right side we identified the image of M by the forgetful functor with M, since it's simply the underlying set.) Even without any category theory, one often describes the free R-module on a set I by the following universal property: given a function f from a set I to an R-module M, there exists a unique R-module homomorphism $\oplus_I R \to M$ that agrees with f on the basis elements. And indeed, this is what is encoded in the bijection above.

This type of *free* construction is quite common, and often gives rise to adjunctions. We can think about the free functor from **Set** to R-**Mod** as the most efficient way of defining an R-module from a given set. It's efficient because it comes with a nice universal property.

Quoting Mac Lane [ML98], one of the fathers of category theory, "the slogan is *adjoint functors arise everywhere*". We will see a very important example of adjunction later on – the Hom-tensor adjunction.

Remark 1.117. We can rephrase the condition that $G\colon \mathscr{D} \longrightarrow \mathscr{C}$ has a left adjoint functor $F : \mathscr{C} \longrightarrow \mathscr{D}$ as follows: for every object C in $\mathscr{C}$, there is a universal arrow from C to G, and for every object D in $\mathscr{D}$ there exists a universal arrow from F to D. To see that, let $\eta_D \in \mathrm{Hom}_{\mathscr{D}}(F(G(D)), D)$ be the image of the identity on $\mathrm{Hom}_{\mathscr{D}}(G(D), G(D))$ via the bijection

$$\begin{array}{ccc} \mathrm{Hom}_{\mathscr{C}}(G(D), G(D)) & \xrightarrow{\cong} & \mathrm{Hom}_{\mathscr{D}}(F(G(D)), D) \\ \mathrm{id}_{G(D)} & \longmapsto & \eta_D \end{array}$$

given by the definition of adjoint functors, and let $\varepsilon_C \in \mathrm{Hom}_{\mathscr{C}}(C, GF(C))$ be be the image of the identity on $\mathrm{Hom}_{\mathscr{C}}(F(C), F(C))$ via the bijection

$$\begin{array}{ccc} \mathrm{Hom}_{\mathscr{D}}(F(C), F(C)) & \xrightarrow{\cong} & \mathrm{Hom}_{\mathscr{D}}(C, GF(C)) \ . \\ \mathrm{id}_{F(C)} & \longmapsto & \varepsilon_C \end{array}$$

We claim that $(F(C), \varepsilon_C)$ is a universal arrow from C to G. That would mean that given arrow $f \in \mathrm{Hom}_{\mathscr{C}}(C, G(Y))$, there must exist a unique arrow $h \in \mathrm{Hom}_{\mathscr{D}}(F(C), Y)$ such that the following diagram commutes:

$$\begin{array}{ccc} F(C) & & D \xrightarrow{\varepsilon_C} G(F(C)) \\ \Big\downarrow{\scriptstyle h} & & \searrow^{f} \Big\downarrow{\scriptstyle G(h)} \\ Y & & G(Y). \end{array}$$

This says that $G(h)_*(\varepsilon_C) = G(h) \circ \varepsilon_C = f$, which means that

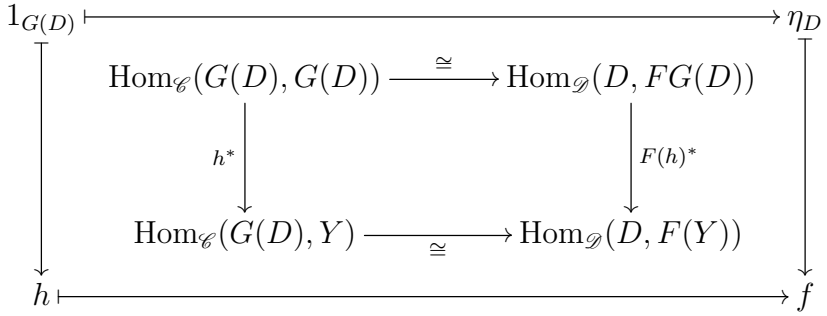
$$\begin{array}{ccc} 1_{F(C)} & \longmapsto & \varepsilon_C \\ \Big\downarrow & & \Big\downarrow \\ \mathrm{Hom}_{\mathscr{D}}(F(C), F(C)) & \xrightarrow{\cong} & \mathrm{Hom}_{\mathscr{C}}(C, GF(C)) \\ \Big\downarrow{\scriptstyle h_*} & & \Big\downarrow{\scriptstyle G(h)_*} \\ \mathrm{Hom}_{\mathscr{D}}(F(C), Y) & \xrightarrow[\cong]{} & \mathrm{Hom}_{\mathscr{C}}(C, G(Y)) \\ \Big\downarrow & & \Big\downarrow \\ h & \longmapsto & f \end{array}$$

On the one hand, such an h does exist: just take $h \in \mathrm{Hom}_{\mathscr{C}}(F(C), Y)$ that is sent to f via the bijection between $\mathrm{Hom}_{\mathscr{D}}(F(C), Y)$ and $\mathrm{Hom}_{\mathscr{C}}(C, G(Y))$. Since this map is a bijection, such an h is unique.

Similarly, we claim that $(G(D), \eta_D)$ is a universal arrow from F to D. That would mean that for any arrow $f \in \mathrm{Hom}_{\mathscr{D}}(F(Y), D)$, there exists a unique $h \in \mathrm{Hom}_{\mathscr{C}}(Y, G(D))$ such that the following diagram commutes:

$$\begin{array}{ccc} G(D) & & D \xleftarrow{\eta_D} F(G(D)) \\ \Big\uparrow{\scriptstyle h} & & \nwarrow^{f} \Big\uparrow{\scriptstyle F(h)} \\ Y & & F(Y) \end{array}$$

This means that $(F(h))^*(\eta_D) = \eta_D \circ F(h) = f$, so



Again, such an h exists and it is unique because it must correspond to f via the bijection between $\mathrm{Hom}_{\mathscr{D}}(D, F(Y))$ and $\mathrm{Hom}_{\mathscr{C}}(G(D), Y)$.

We can talk about *the* left or right adjoint to a given functor.

Exercise 1.118. Left and right adjoints are unique up to natural isomorphism. More precisely, given an adjoint pair of functors (F, G), show that if G' is also a right adjoint to F, then G' and G are naturally isomorphic. Similarly, show that if F' is also a left adjoint to G, then F and F' are naturally isomorphic.

We close this short detour into the wonderful world of category theory to point out that if we wanted to sound really obscure, we could have defined chain complexes in this categorical language.

Remark 1.119. First, we view $\mathbb{Z}$ as a partially ordered set under $\geqslant$. As in Theorem 1.9, $\mathbb{Z}$ now gives us a category whose objects are the integers, and where we have an arrow in $\mathrm{Hom}_{\mathbb{Z}}(n, m)$ if $n \geqslant m$. If we ignore the identity maps $\mathrm{Hom}_{\mathbb{Z}}(n, n)$ and composite maps, we can represent this category in the following diagram:

$$\cdots \longrightarrow n+1 \longrightarrow n \longrightarrow n-1 \longrightarrow \cdots .$$

From this perspective, a chain complex is a functor $F\colon \mathbb{Z} \longrightarrow R\text{-}\mathbf{Mod}$: for each $n \in \mathbb{Z}$, we get an R-module F_n, and we also get an R-module homomorphisms $F_{n+1} \longrightarrow F_n$ for each n. Indeed, this can all be represented as a sequence

$$\cdots \longrightarrow F_{n+1} \longrightarrow F_n \longrightarrow F_{n-1} \longrightarrow \cdots .$$

For our functor to truly be a complex, though, we must require that all compositions $F_{n+1} \longrightarrow F_n \longrightarrow F_{n-1}$ be 0. A map of complexes, also known as a chain map, is a natural transformation between two such functors.

Chapter 2

The category of chain complexes

We are finally ready to introduce the category of chain complexes, and to talk more about exact sequences and homology.

2.1 Maps of complexes

Unsurprisingly, we can form a category of complexes, but to do that we need the right definition of maps between complexes. We also take this section as a chance to set up some definitions we will need later. One thing to keep in mind as we build our basic definitions: we also want homology to be functorial.

Definition 2.1. Let $(F_\bullet, \partial^F_\bullet)$ and $(G_\bullet, \partial^G_\bullet)$ be complexes. A **map of complexes** or a **chain map**, which we write as $h : (F_\bullet, \partial^F_\bullet) \longrightarrow (G_\bullet, \partial^G_\bullet)$ or simply $h : F \longrightarrow G$, is a sequence of homomorphisms of R-modules $h_n \colon F_n \longrightarrow G_n$ such that the following diagram commutes:

$$\begin{array}{ccccccccc} \cdots & \longrightarrow & F_{n+1} & \longrightarrow & F_n & \longrightarrow & F_{n-1} & \longrightarrow & \cdots \\ & & \downarrow{\scriptstyle h_{n+1}} & & \downarrow{\scriptstyle h_n} & & \downarrow{\scriptstyle h_{n-1}} & & \\ \cdots & \longrightarrow & G_{n+1} & \longrightarrow & G_n & \longrightarrow & G_{n-1} & \longrightarrow & \cdots \end{array}$$

This means that $h_n \partial^F_{n+1} = \partial^G_{n+1} h_{n+1}$ for all n.

Note that throughout, whenever we call a function $f : M \to N$ between R-modules M and N a *map*, we really mean to say it is a homomorphism of R-modules.

Example 2.2. The zero and the identity maps of complexes $(F_\bullet, \partial_\bullet) \longrightarrow (F_\bullet, \partial_\bullet)$ are exactly what they sound like: the zero map $0_{F_\bullet}$ is 0 in every homological degree, and the identity map $1_{F_\bullet}$ is the identity in every homological degree.

This is the notion of morphism we would want to form a category of chain complexes.

Definition 2.3. Let R be a ring. The **category of chain complexes** of R-modules, denoted $\mathrm{Ch}(R\text{-}\mathbf{mod})$ or simply $\mathrm{Ch}(R)$, is the category with objects all chain complexes of R-modules and arrows all maps of complexes of R-modules. When $R = \mathbb{Z}$, we write $\mathrm{Ch}(\mathbf{Ab})$ for $\mathrm{Ch}(\mathbb{Z})$, the category of chain complexes of abelian groups.

Note that the identity maps defined above are precisely the identity arrows in the category of chain complexes.

Exercise 2.4. Show that the isomorphisms in the category $\mathrm{Ch}(R)$ are precisely the maps of complexes

$$\begin{array}{ccccccccc} \cdots & \longrightarrow & F_{n+1} & \longrightarrow & F_n & \longrightarrow & F_{n-1} & \longrightarrow & \cdots \\ & & \downarrow{\scriptstyle h_{n+1}} & & \downarrow{\scriptstyle h_n} & & \downarrow{\scriptstyle h_{n-1}} & & \\ \cdots & \longrightarrow & G_{n+1} & \longrightarrow & G_n & \longrightarrow & G_{n-1} & \longrightarrow & \cdots \end{array}$$

such that h_n is an isomorphism for all n.

This is a good notion of map of complexes: it induces homomorphisms in homology, which in particular allows us to say that homology is a functor.

Lemma 2.5. *Let $h : (F_\bullet, \partial^F_\bullet) \longrightarrow (G_\bullet, \partial^G_\bullet)$ be a map of complexes. For all n, h_n restricts to homomorphisms $B_n(h) : B_n(F_\bullet) \longrightarrow B_n(G_\bullet)$ and $Z_n(h) : Z_n(F_\bullet) \longrightarrow Z_n(G_\bullet)$. As a consequence, h induces homomorphisms on homology $\mathrm{H}_n(h) : \mathrm{H}_n(F_\bullet) \longrightarrow \mathrm{H}_n(G_\bullet)$.*

Proof. Since $h_n\partial^F_{n+1} = \partial^G_{n+1}h_{n+1}$, any element $a \in B_n(F_\bullet)$, say $a = \partial^F_{n+1}(b)$, is taken to

$$h_n(a) = h_n\partial^F_{n+1}(b) = \partial^G_{n+1}h_{n+1}(b) \in \operatorname{im}\partial^G_{n+1} = B_n(G_\bullet).$$

Similarly, if $a \in Z_n(F_\bullet) = \ker\partial^F_n$, then

$$\partial_n h_n(a) = h_{n-1}\partial^F_n(a) = 0,$$

so $h_n(a) \in \ker\partial^G_n = Z_n(G_\bullet)$. Finally, the restriction of h_n to $Z_n(F_\bullet) \longrightarrow Z_n(G_\bullet)$ sends $B_n(F_\bullet)$ into $B_n(G_\bullet)$, and thus it induces a well-defined homomorphism on the quotients $\mathrm{H}_n(F_\bullet) \longrightarrow \mathrm{H}_n(G_\bullet)$. □

Definition 2.6. Let $h\colon (F_\bullet, \partial^F_\bullet) \longrightarrow (G_\bullet, \partial^G_\bullet)$ be a map of complexes. We call the map

$$\begin{aligned} \mathrm{H}_n(h)\colon \quad \mathrm{H}_n(F_\bullet) &\longrightarrow \mathrm{H}_n(G_\bullet) \\ a + B_n(F) &\mapsto h_n(a) + B_n(G) \end{aligned}$$

the **induced map in homology**, and sometimes denote it by h_*.

One can show that H_n preserves compositions, and that moreover, the map in homology induced by the identity is the identity. Thus taking nth homology is a functor

$$\mathrm{H}_n\colon \mathrm{Ch}(R) \longrightarrow R\textbf{-Mod}$$

which takes each map of complexes $h\colon F_\bullet, \longrightarrow G_\bullet$ to the R-module homomorphism

$$\mathrm{H}_n(h) : \mathrm{H}_n(F_\bullet) \longrightarrow \mathrm{H}_n(G_\bullet).$$

Definition 2.7. A map of chain complexes h is a **quasi-isomorphism** if it induces an isomorphism in homology, meaning $\mathrm{H}_n(h)$ is an isomorphism of R-modules for all n. If there exists a quasi-isomorphism between two complexes C and D, we say that C and D are **quasi-isomorphic**, and write $C \simeq D$.

Remark 2.8. Note that saying that if f is a quasi-isomorphism between F and G is a stronger statement that the fact that $\mathrm{H}_n(F) \cong \mathrm{H}_n(G)$ for all n: it also says that there are isomorphisms $\mathrm{H}_n(F) \cong \mathrm{H}_n(G)$ that are all induced by f.

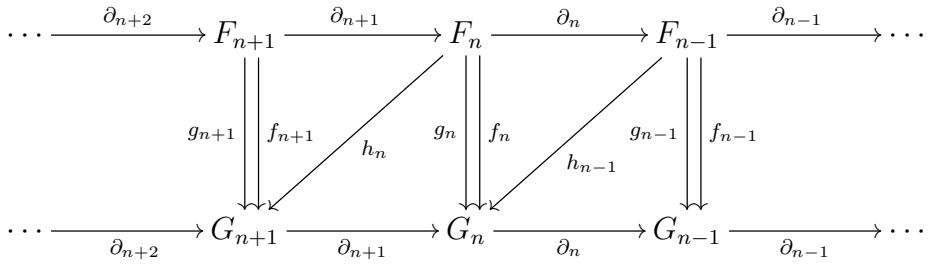
Not all quasi-isomorphisms are isomorphisms, as the following example shows:

Exercise 2.9. Let π denote the projection map from $\mathbb{Z}$ to $\mathbb{Z}/2\mathbb{Z}$. The chain map

$$\begin{array}{ccccccccccc}
\cdots & \longrightarrow & 0 & \longrightarrow & \mathbb{Z} & \xrightarrow{2} & \mathbb{Z} & \longrightarrow & 0 & \longrightarrow & \cdots \\
 & & \downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle 0} & & \\
\cdots & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \mathbb{Z}/2\mathbb{Z} & \longrightarrow & 0 & \longrightarrow & \cdots
\end{array}$$

is a quasi-isomorphism.

Definition 2.10. Let $f, g\colon F \longrightarrow G$ be maps complexes. A **homotopy**, sometimes referred to as a **chain homotopy**, between f and g is a sequence of maps $h_n\colon F_n \longrightarrow G_{n+1}$



such that

$$\partial_{n+1} h_n + h_{n-1} \partial_n = f_n - g_n$$

for all n. If there exists a homotopy between f and g, we say that f and g are **homotopic** or that they **have the same homotopy type**. We write $f \simeq g$ to say that f and g are homotopic. If f is homotopic to the zero map, we say f is **nullhomotopic**, and write $f \simeq 0$. This should not be confused with the notation $C \simeq D$ on complexes.

Exercise 2.11. Homotopy is an equivalence relation.

The equivalence classes under homotopy are called **homotopy classes**. Homotopy is an interesting equivalence relation because homotopic maps induce the same map on homology.

Lemma 2.12. *Let $f, g\colon (F_\bullet, \partial^F_\bullet) \longrightarrow (G_\bullet, \partial^G_\bullet)$ be maps of complexes. If f is homotopic to g, then $\mathrm{H}_n(f) = \mathrm{H}_n(g)$ for all n. In particular, every nullhomotopic map induces the zero map in homology.*

Proof. Let $f, g : (F_\bullet, \partial^F_\bullet) \longrightarrow (G_\bullet, \partial^G_\bullet)$ be homotopic maps of complexes, and let h be a homotopy between f and g. We claim that the map of complexes $f - g$ (defined in the obvious way) sends cycles to boundaries. If $a \in Z_n(F_\bullet)$, then

$$(f-g)_n(a) = \partial_{n+1} h_n(a) + h_{n-1} \underbrace{\partial_n(a)}_{0} = \partial_{n+1}(h_n(a)) \in B_n(G_\bullet).$$

The map on homology induced by $f - g$ must then be the 0 map, so f and g induce the same map on homology. Here we are implicitly using the fact that $\mathrm{H}_n(f+h) = \mathrm{H}_n(f) + \mathrm{H}_n(g)$, which we leave as an exercise to be further explored in Theorem 3.9. □

Notice, however, that the converse is false: the induced map in homology can be the zero map (for all homological degrees) even if the original map of complexes is not nullhomotopic.

Exercise 2.13. Consider the following map of complexes:

$$\begin{array}{ccccccccccc}
\cdots & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \mathbb{Z}/2 & \longrightarrow & 0 & \longrightarrow & \cdots \\
 & & \downarrow & & \downarrow & & \downarrow{\scriptstyle 2} & & \downarrow & & \\
\cdots & \longrightarrow & 0 & \longrightarrow & \mathbb{Z}/4\mathbb{Z} & \xrightarrow[2]{} & \mathbb{Z}/4\mathbb{Z} & \longrightarrow & 0 & \longrightarrow & \cdots
\end{array}$$

Show that this map is not nullhomotopic, but that the induced map in homology is zero.

Definition 2.14. If $f : (F_\bullet, \partial^F_\bullet) \longrightarrow (G_\bullet, \partial^G_\bullet)$ and $g : (G_\bullet, \partial^G_\bullet) \longrightarrow (F_\bullet, \partial^F_\bullet)$ are maps of complexes such that fg is homotopic to the identity map on $(G_\bullet, \partial^G_\bullet)$ and gf is homotopic to the identity chain map on $(F_\bullet, \partial^F_\bullet)$, we say that f and g are **homotopy equivalences** and $(F_\bullet, \partial^F_\bullet)$ and $(G_\bullet, \partial^G_\bullet)$ are **homotopy equivalent**.

Corollary 2.15. *Homotopy equivalences are quasi-isomorphisms.*

Proof. If $f\colon (F_\bullet, \partial^F_\bullet) \longrightarrow (G_\bullet, \partial^G_\bullet)$ and $g\colon (G_\bullet, \partial^G_\bullet) \longrightarrow (F_\bullet, \partial^F_\bullet)$ are such that fg is homotopic to $1_{G_\bullet}$ and gf is homotopic to $1_{F_\bullet}$, then by Theorem 2.12 the map fg induces the identity map on homology. So for all n we have

$$\mathrm{H}_n(f)\,\mathrm{H}_n(g) = \mathrm{H}_n(fg) = \mathrm{H}_n(1) = 1.$$

Therefore, $\mathrm{H}_n(f)$ and $\mathrm{H}_n(g)$ must both be isomorphisms. □

The converse is false.

Exercise 2.16. Let π denote the projection map from $\mathbb{Z}$ to $\mathbb{Z}/2\mathbb{Z}$. The chain map

$$\begin{array}{ccccccccccc}
\cdots & \longrightarrow & 0 & \longrightarrow & \mathbb{Z} & \xrightarrow{2} & \mathbb{Z} & \longrightarrow & 0 & \longrightarrow & \cdots \\
 & & \downarrow & & \downarrow & & \downarrow{\scriptstyle \pi} & & \downarrow & & \\
\cdots & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \mathbb{Z}/2\mathbb{Z} & \longrightarrow & 0 & \longrightarrow & \cdots
\end{array}$$

is a quasi-isomorphism but not a homotopy equivalence.

Remark 2.17. The relation $F \simeq G$, meaning "there is a quasi-isomorphism from F to G", is not symmetric: in Theorem 2.16, there is no quasi-isomorphism going in the opposite direction of the one given.

Now that we know about maps between complexes, it's time to point out that we can also talk about complexes of complexes and exact sequences of complexes. While we will later formalize this a little better when we discover that $\mathrm{Ch}(R)$ is an abelian category, let's for now give quick definitions that we can use.

Definition 2.18. Given complexes B and C, B is a **subcomplex** of C if B_n is a submodule of C_n for all n, and the inclusion maps $\iota_n\colon B_n \subseteq C_n$ define a map of complexes $\iota : B \longrightarrow C$. Given a subcomplex B of C, the **quotient** of C by B is the complex C/B that has C_n/B_n in homological degree n, with differential induced by the differential on C_n.

Exercise 2.19. If B is a subcomplex of C, then the differential d on C satisfies $d_n(B_n) \subseteq B_{n-1}$. Therefore, d_n induces a map of R-modules $C_n/B_n \longrightarrow C_{n-1}/B_{n-1}$ for all n, so that our definition of the differential on C/B actually makes sense.

We can also talk about kernels and cokernels of maps of complexes.

Definition 2.20. Given any map of complexes $f\colon B_\bullet \longrightarrow C_\bullet$, the **kernel** of f is the subcomplex $\ker f$ of $B_\bullet$ that we can assemble from the the kernels $\ker f_n$. More precisely, $\ker f$ is the complex

$$\cdots \longrightarrow \ker f_{n+1} \longrightarrow \ker f_n \longrightarrow \ker f_{n-1} \longrightarrow \cdots$$

where the differentials are simply the corresponding restrictions of the differentials on $B_\bullet$. Similarly, the **image** of f is the subcomplex of $C_\bullet$

$$\cdots \longrightarrow \operatorname{im} f_{n+1} \longrightarrow \operatorname{im} f_n \longrightarrow \operatorname{im} f_{n-1} \longrightarrow \cdots$$

where the differentials are given by restriction of the corresponding differentials in $C_\bullet$. The **cokernel** of f is the quotient complex $C_\bullet/\operatorname{im} f$.

Again, there are some details to check.

Exercise 2.21. Show that the kernel, image, and cokernel of a complex map are indeed complexes.

Definition 2.22. A **complex** in $\mathrm{Ch}(R)$ is a sequence of complexes of R-modules C_n and chain maps $d_n\colon C_n \longrightarrow C_{n-1}$ between them

$$\cdots \longrightarrow C_{n+1} \xrightarrow{d_{n+1}} C_n \xrightarrow{d_n} C_{n-1} \longrightarrow \cdots$$

such that $d_n d_{n+1} = 0$ for all n. A complex of complexes is a diagram of the form

$$\begin{array}{ccccccc}
\cdots \longrightarrow & C_{n+1,i+1} & \xrightarrow{d_{n+1}} & C_{n,i+1} & \xrightarrow{d_n} & C_{n-1,i+1} & \longrightarrow \cdots \\
& \downarrow{\scriptstyle \partial_{i+1}} & & \downarrow{\scriptstyle \partial_{i+1}} & & \downarrow{\scriptstyle \partial_{i+1}} & \\
\cdots \longrightarrow & C_{n+1,i} & \xrightarrow{d_{n+1}} & C_{n,i} & \xrightarrow{d_n} & C_{n-1,i} & \longrightarrow \cdots \\
& \downarrow{\scriptstyle \partial_i} & & \downarrow{\scriptstyle \partial_i} & & \downarrow{\scriptstyle \partial_i} & \\
\cdots \longrightarrow & C_{n+1,i-1} & \xrightarrow{d_{n+1}} & C_{n,i-1} & \xrightarrow{d_n} & C_{n-1,i-1} & \longrightarrow \cdots
\end{array}$$

where $C_{i,j}$ is the module in homological degree j in the complex C_i. The nth column corresponds to the complex C_n, and every row is also a complex. The vertical maps are the differentials on each individual complex; the horizontal maps are the differentials on the complex of complexes.

Given a complex C in $\mathrm{Ch}(R)$, we can talk about cycles and boundaries, which are a sequence of subcomplexes of the complexes in C, and thus its homology. Such a complex is exact if $\operatorname{im} d_{n+1} = \ker d_n$ for all n.

Definition 2.23. A **short exact sequence** of complexes is an exact complex in $\mathrm{Ch}(R)$ of the form

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0.$$

Equivalently, a short exact sequence of complexes is a commutative diagram

$$\begin{array}{ccccccccc}
 & & \vdots & & \vdots & & \vdots & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & A_{i+1} & \xrightarrow{f_{i+1}} & B_{i+1} & \xrightarrow{g_{i+1}} & C_{i+1} & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \partial_{i+1}} & & \downarrow{\scriptstyle \partial_{i+1}} & & \downarrow{\scriptstyle \partial_{i+1}} & & \\
0 & \longrightarrow & A_i & \xrightarrow{f_i} & B_i & \xrightarrow{g_i} & C_i & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \partial_i} & & \downarrow{\scriptstyle \partial_i} & & \downarrow{\scriptstyle \partial_i} & & \\
0 & \longrightarrow & A_{i-1} & \xrightarrow{f_{i-1}} & B_{i-1} & \xrightarrow{g_{i-1}} & C_{i-1} & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & \vdots & & \vdots & & \vdots & &
\end{array}$$

where the rows are exact and the columns are complexes.

2.2 Short exact sequences

In this section, we will discuss short exact sequences of modules in a bit more detail. We note, however, that everything we will discuss here can be extended for short exact sequences of complexes, and that the generalization is not too difficult: one just needs to replace modules with complexes and maps of modules by maps of complexes.

Example 2.24. Fix a ring R, and let A and C be R-modules. Consider the inclusion $i : A \to A \oplus C$ of A into the first component of the direct sum, and the projection map $\pi : A \oplus C \to C$ onto the second component of the product. These two maps fit into a short exact sequence

$$0 \longrightarrow A \xrightarrow{i} A \oplus C \xrightarrow{p} C \longrightarrow 0.$$

These are sometimes called **trivial short exact sequences**.

On the one hand, the short exact sequences that look like this one are very important; on the other hand, not all short exact sequences are of this type.

Definition 2.25. We say that a short exact sequence

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

splits or is a **split short exact sequence** if it is isomorphic to

$$0 \longrightarrow A \xrightarrow{i} A \oplus C \xrightarrow{p} C \longrightarrow 0$$

where i is the inclusion of the first component and p is the projection onto the second component.

Lemma 2.26 (Splitting Lemma)**.** *Consider the short exact sequence*

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

of R-modules. The following are equivalent:

a) There exists a homomorphism of R-modules $q\colon B \longrightarrow A$ such that $qf = \mathrm{id}_A$.

b) There exists a homomorphism of R-modules $r\colon C \longrightarrow B$ such that $gr = \mathrm{id}_C$.

c) The short exact sequence splits.

Definition 2.27. Given a split short exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0,$$

maps q and r satisfying the conditions of the Splitting Lemma are called **splittings**.

Proof. First, we will show that c) implies a) and b). If the sequence splits, then consider an isomorphism of complexes

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle a} & & \downarrow{\scriptstyle b} & & \downarrow{\scriptstyle c} & & \\
0 & \longrightarrow & A & \xrightarrow[\pi]{i} & A \oplus C & \xrightarrow[j]{p} & C & \longrightarrow & 0,
\end{array}$$

meaning that the diagram commutes and a, b, and c are isomorphisms of R-modules, i is the inclusion in the first component, and p is the projection onto the second component. Let $\pi : A \oplus C \longrightarrow A$ be the projection onto the first component, and $j\colon C \longrightarrow A \oplus C$ be the inclusion into the second component. Now consider the maps $q := a^{-1}\pi b$ and $r := b^{-1}jc$. Then

$$\begin{aligned}
qf &= a^{-1}\pi b f \\
&= a^{-1}\pi i a \quad \text{by commutativity} \\
&= a^{-1}a \qquad \text{because } \pi i = \mathrm{id}_A \\
&= 1_A
\end{aligned}$$

and

$$\begin{aligned}
gr &= gb^{-1}jc \\
&= c^{-1}(cg)b^{-1}jc && \text{multiplying by } c^{-1}c = 1_C \\
&= c^{-1}(pb)b^{-1}jc && \text{by commutativity} \\
&= c^{-1}pjc && \text{because } bb^{-1} = 1_B \\
&= c^{-1}c && \text{because } pj = \mathrm{id}_C \\
&= 1_C.
\end{aligned}$$

Therefore, c) implies a) and b).

Now suppose that a) holds, and let's show that the sequence splits. First, we need to show that $B \cong A \oplus C$. Every $b \in B$ can be written as

$$b = (b - fq(b)) + fq(b),$$

where $fq(b) \in \operatorname{im} f \cong A$, and

$$q(b - fq(b)) = q(b) - \underbrace{qf}_{\mathrm{id}_A}(q(b)) = q(b) - q(b) = 0,$$

so $b - fq(b) \in \ker q$. This shows that $B = \operatorname{im} f + \ker q$. Moreover, if $f(a) \in \ker q$, then $a = qf(a) = 0$, so $\operatorname{im} f \cap \ker q = 0$, and $B = \operatorname{im} f \oplus \ker q$. Now when we restrict g to $\ker q$, g becomes injective. We claim it is also surjective, and thus an isomorphism. Indeed, for any $c \in C$ we can pick $b \in B$ such that $g(b) = c$, since g is surjective, and we showed that we can write $b = f(a) + k$ for some $k \in \ker q$. Then

$$g(k) = \underbrace{gf}_{0}(a) + g(k) = g(b) = c.$$

Finally, note that $\operatorname{im} f \cong A$, so we conclude that $B \cong A \oplus C$, via the isomorphism φ given by

$$\begin{aligned} B &\longrightarrow \operatorname{im} f \oplus \ker q \longrightarrow A \oplus C \\ b &\longmapsto (fq(b), b - fq(b)) \longmapsto (q(b), g(b)). \end{aligned}$$

Since $gf = 0$ and $qf = \mathrm{id}_A$, $\varphi f(a) = (qf(a), 0) = (a, 0)$, so $\varphi f = i$, where $i \colon A \longrightarrow A \oplus C$ is the inclusion on the first factor. If $p \colon A \oplus C \longrightarrow C$ denotes the projection onto the second factor, $p\varphi = g$. Together, these two facts say that the following is a map of complexes:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \longrightarrow & 0 \\ & & \| & & \downarrow{\scriptstyle \varphi} & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow[i]{} & A \oplus C & \xrightarrow[p]{} & C & \longrightarrow & 0. \end{array}$$

Since φ is an isomorphism, so is our map of complexes, and thus our original sequence is a split exact sequence. This shows that a) implies c).

Now assume b) holds. Every $b \in B$ can be written as

$$b = (b - rg(b)) + rg(b),$$

where $rg(b) \in \operatorname{im} r$ and

$$g(b - rg(b)) = g(b) - \underbrace{gr}_{\mathrm{id}_C}(g(b)) = g(b) - g(b) = 0,$$

so $b - rg(b) \in \ker g$. This shows that $B = \ker g + \operatorname{im} r$. Moreover, if $r(c) \in \ker g$, then

$$c = \mathrm{id}_C(c) = gr(c) = 0.$$

Therefore, $B = \ker g \oplus \operatorname{im} r$. Now r is injective, since $r(c) = 0 \implies c = gr(c) = 0$, and thus $\operatorname{im} r \cong C$. Since $\ker g = \operatorname{im} f \cong A$, we conclude that $B \cong A \oplus C$, via the isomorphism

$$\begin{aligned} A \oplus C &\xrightarrow{\quad\psi\quad} B \\ (a, c) &\longmapsto f(a) + r(c). \end{aligned}$$

Finally, let $i \colon A \longrightarrow A \oplus C$ denote the inclusion of the first factor, and $p \colon A \oplus C \longrightarrow C$ denote the projection onto the second factor. By construction, $\psi i = f$. Moreover,

$$g\psi(a, c) = \underbrace{gf}_{0}(a) + \underbrace{gr}_{\mathrm{id}_C}(c) = c,$$

so $g\psi = p$. Together, these say that the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{i} & A \oplus C & \xrightarrow{p} & C & \longrightarrow & 0 \\ & & \| & & \downarrow{\scriptstyle \psi} & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow[f]{} & B & \xrightarrow[g]{} & C & \longrightarrow & 0 \end{array}$$

commutes, and must then be an isomorphism of short exact sequences. □

Remark 2.28. In the split short exact sequence

$$0 \longrightarrow A \xrightarrow{i} A \oplus C \xrightarrow{p} C \longrightarrow 0,$$

the canonical projection $q \colon A \oplus C \to A$ and the usual inclusion $r \colon C \to A \oplus C$ are splittings.

Exercise 2.29. Let k be a field. Show that every short exact sequence of k-vector spaces splits.

The Rank-Nulity Theorem can be recast in this setting as a consequence of the fact that every short exact sequence of k-vector spaces splits.

Exercise 2.30. Prove the Rank-Nulity Theorem using Theorem 2.29: show that given any linear transformation $T \colon V \to W$ of k-vector spaces,

$$\dim(\operatorname{im} T) + \dim(\ker T) = \dim V.$$

But over a general ring, not every short exact sequence splits.

Example 2.31. The short exact sequence

$$0 \longrightarrow \mathbb{Z} \xrightarrow{2} \mathbb{Z} \longrightarrow \mathbb{Z}/2 \longrightarrow 0$$

is not split. Indeed, $\mathbb{Z}$ does not have any 2-torsion elements, so it is not isomorphic to $\mathbb{Z} \oplus \mathbb{Z}/2$.

An alternative explanation is that there is no splitting to the inclusion $\mathbb{Z} \xrightarrow{2} \mathbb{Z}$. On the one hand, every $\mathbb{Z}$-module map is given by multiplication by a fixed integer n, so a splitting $f : \mathbb{Z} \longrightarrow \mathbb{Z}$ would be of the form $f(a) = na$ for some fixed n. On the other hand, our proposed splitting f must send 2 to 1, but there is no integer solution n to $2n = f(2) = 1$.

More surprisingly, a short exact sequence of the form

$$0 \longrightarrow A \xrightarrow{f} A \oplus C \xrightarrow{g} C \longrightarrow 0$$

is not necessarily split, not unless f is the inclusion of the first component and g is the projection onto the second component, as the next example will show.

Example 2.32. Consider the short exact sequence

$$0 \longrightarrow \mathbb{Z}/(2) \xrightarrow{f} \mathbb{Z}/(4) \xrightarrow{g} \mathbb{Z}/(2) \longrightarrow 0$$

where f is the inclusion of the subgroup generated by 2, so $f(1+(2)) = 2+(4)$, and g is the quotient onto that subgroup, meaning $g(1) = 1$. This is not a split short exact sequence, because $\mathbb{Z}/(4) \ncong \mathbb{Z}/(2) \oplus \mathbb{Z}/(2)$. Now let

$$M := \bigoplus_{\mathbb{N}} (\mathbb{Z}/(2) \oplus \mathbb{Z}/(4))$$

be the direct sum of infinitely many copies of $\mathbb{Z}/(2) \oplus \mathbb{Z}/(4)$. Then

$$\mathbb{Z}/(2) \oplus M \cong M \cong M \oplus \mathbb{Z}/(4),$$

and the sequence

$$0 \longrightarrow \mathbb{Z}/(2) \xrightarrow{h} \mathbb{Z}/(4) \oplus M \xrightarrow{t} \mathbb{Z}/(2) \oplus M \longrightarrow 0$$

with $h(a) = (f(a), 0)$ and $t(a, m) = (g(a), m)$ is still exact. The middle term is indeed isomorphic to the direct sum of the other two:

$$\mathbb{Z}/(4) \oplus M \cong M \cong (M \oplus \mathbb{Z}/(2)) \oplus \mathbb{Z}/(2).$$

And yet this is not a split exact sequence: if we had a splitting $q : \mathbb{Z}/(4) \oplus M \longrightarrow \mathbb{Z}/(2)$ of h, then its restriction to the first factor would give us a splitting $\mathbb{Z}/(4) \longrightarrow \mathbb{Z}/(2)$ of f, which we know cannot exist, since

$$0 \longrightarrow \mathbb{Z}/(2) \xrightarrow{f} \mathbb{Z}/(4) \xrightarrow{g} \mathbb{Z}/(2) \longrightarrow 0$$

does not split.

Given splittings q and r for a short exact sequence as in Theorem 2.26, we can quickly show that our short exact sequence splits using the Five Lemma. To prove the Five Lemma, one needs to use diagram chasing. Diagram chasing is a common technique in homological algebra, which essentially consists of tracing elements around in the diagram. We will see some examples of diagram chasing in the next section.

Exercise 2.33 (The Five Lemma)**.** Consider the following commutative diagram of R-modules with exact rows:

$$\begin{array}{ccccccccc} A' & \longrightarrow & B' & \longrightarrow & C' & \longrightarrow & D' & \longrightarrow & E' \\ \downarrow{\scriptstyle a} & & \downarrow{\scriptstyle b} & & \downarrow{\scriptstyle c} & & \downarrow{\scriptstyle d} & & \downarrow{\scriptstyle e} \\ A & \longrightarrow & B & \longrightarrow & C & \longrightarrow & D & \longrightarrow & E \end{array}$$

Show that if a, b, d, and e are isomorphisms, then c is an isomorphism.

Remark 2.34. Given a short exact sequence, suppose we have R-module homomorphisms q and r

$$0 \longrightarrow A \underset{q}{\overset{f}{\rightleftarrows}} B \underset{r}{\overset{g}{\rightleftarrows}} C \longrightarrow 0$$

such that $qf = \mathrm{id}_A$ and $rg = \mathrm{id}_C$. Then we get an induced map

$$\begin{aligned} B &\xrightarrow{\varphi} A \oplus C \\ b &\longmapsto (q(b), g(b)) \end{aligned}$$

such that the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \longrightarrow & 0 \\ & & \| & & \downarrow{\scriptstyle \varphi} & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow[i]{} & A \oplus C & \xrightarrow[p]{} & C & \longrightarrow & 0. \end{array}$$

commutes. The Five Lemma guarantees that φ must be an isomorphism, so our diagram is an isomorphism of short exact sequences.

There are many ways in which R-**Mod** behaves better than the category of groups, and this is one of them.

Remark 2.35. The Splitting Lemma does not hold if we replace R-modules with the category **Grp** of groups. For example, consider the symmetric group on 3 elements S_3 and the inclusion $A_3 \hookrightarrow S_3$ of the alternating group in S_3. Notice that A_3 is precisely the kernel of the sign map

$$\mathrm{sign}\colon S_3 \longrightarrow \mathbb{Z}/2,$$

which sends even permutations to 0 and odd permutations to 1. Therefore,

$$0 \longrightarrow A_3 \longrightarrow S_3 \longrightarrow \mathbb{Z}/2 \longrightarrow 0$$

is a short exact sequence. When writing exact sequences of nonabelian groups such as this one, one sometimes uses $\{e\}$ instead of 0, to indicate that trivial group. So our short exact sequence is

$$\{e\} \longrightarrow A_3 \longrightarrow S_3 \longrightarrow \mathbb{Z}/2 \longrightarrow \{e\}.$$

Moreover, this exact sequence is not split, since S_3 is not abelian but $A_3 \oplus \mathbb{Z}/2$ is, and thus $S_3 \ncong A_3 \oplus \mathbb{Z}/2$. However, any group homomorphism $u\colon \mathbb{Z}/2 \to S_3$ defined by sending the generator to any two cycle is a splitting for our short exact sequence, meaning $\mathrm{sign} \circ u = \mathrm{id}_{\mathbb{Z}/2}$.

Funny enough, there is no splitting for the inclusion $A_3 \subseteq S_3$, since there are no nontrivial homomorphisms $S_3 \to A_3$: A_3 has no elements of order 2, so a group homomorphism $S_3 \to A_3$ must send every 2-cycle in S_3 must be sent to the identity, but 2-cycles generate S_3.

We will return to the topic of split short exact sequences when we talk about projective and injective modules.

Exercise 2.36. Fix a ring R. Show that if F is a free R-module, then every short exact sequence of R-modules

$$0 \longrightarrow A \longrightarrow B \longrightarrow F \longrightarrow 0$$

splits.

2.3 Long exact sequences

A long exact sequence is just what it sounds like: an exact sequence that is, well, long. Usually, we use the term long exact sequence to refer to any exact sequence, especially if it is not a short exact sequence. So in particular, a long exact sequence does not literally have to be that long.

Long exact sequences arise naturally in various ways, and are often induced by some short exact sequence. The first long exact sequence one encounters is the long exact sequence on homology. All other long exact sequences are, in some way, a special case of this one. The main tool we need to build it is the Snake Lemma.

Theorem 2.37 (Snake Lemma)**.** *Consider the commutative diagram of R-modules*

$$\begin{array}{ccccccccc} & & A' & \xrightarrow{i'} & B' & \xrightarrow{p'} & C' & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle h} & & \\ 0 & \longrightarrow & A & \xrightarrow[i]{} & B & \xrightarrow[p]{} & C & & \end{array} .$$

Suppose that the rows of the diagram are exact. Given $c' \in \ker h$, pick $b' \in B'$ such that $p'(b') = c'$, and $a \in A$ such that $i(a) = g(b')$, and consider the map $\partial \colon \ker h \to \operatorname{coker} f$ given by

$$\partial(c') = a + \operatorname{im} f \in \operatorname{coker} f.$$

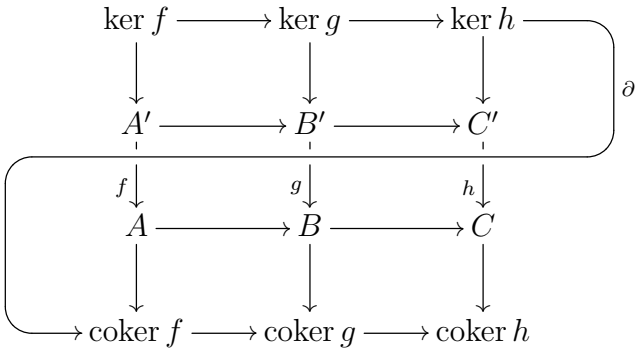
Then there exists an exact sequence

$$\ker f \longrightarrow \ker g \longrightarrow \ker h \xrightarrow{\partial} \operatorname{coker} f \longrightarrow \operatorname{coker} g \longrightarrow \operatorname{coker} h$$

Moreover:

- *If i' is injective then we can extend our exact sequence to $0 \longrightarrow \ker f \longrightarrow \ker g$.*
- *If p is surjective then we can extend our exact sequence to $\operatorname{coker} g \longrightarrow \operatorname{coker} h \longrightarrow 0$.*

The picture to keep in mind (and which explains the name of the lemma) is the following:



Definition 2.38. The map ∂ in the Snake Lemma is the **connecting homomorphism**.

Proof. If $a' \in \ker f$, then

$$g(i'(a')) = if(a') = 0,$$

by commutativity, so $i'(a') \in \ker g$. Similarly, if $b' \in \ker g$ then $p'(b') \in \ker(g)$. So

$$A' \xrightarrow{i'} B' \xrightarrow{p'} C' \qquad \text{restrict to maps} \qquad \ker f \xrightarrow{i'} \ker g \xrightarrow{p'} \ker h\ .$$

We claim that the sequence obtained by restriction

$$\ker f \xrightarrow{i'} \ker g \xrightarrow{p'} \ker h$$

is exact. On the one hand, we already know that the original maps satisfy $p'i' = 0$, so their restrictions must satisfy this as well, guaranteeing that

$$i'(\ker f) \subseteq \ker(\ker g \xrightarrow{p'} \ker h).$$

On the other and, if $b' \in \ker g$ is such that $p'(b') = 0$, then by exactness of the original sequence there exists $a' \in A'$ such that $i'(a') = b'$; we only need to check that we can choose such a' satisfying $a' \in \ker f$. And indeed, by commutativity, any a' with $i'(a') = b'$ satisfies

$$if(a') = gi'(a') = g(b') = 0,$$

and since i is injective, we must have $f(a') = 0$. So we have shown that the following is an exact sequence:

$$\ker f \xrightarrow{i'} \ker g \xrightarrow{p'} \ker h.$$

Similarly, if $a \in \operatorname{im} f$, the commutativity of the diagram guarantees that $i(a) \in \operatorname{im} g$, and if $b \in \operatorname{im} g$, then $p(b) \in \operatorname{im} h$. So the maps $A \xrightarrow{i} B \xrightarrow{p} C$ restrict to maps

$$\operatorname{im} f \xrightarrow{i} \operatorname{im} g \xrightarrow{p} \operatorname{im} h\ ,$$

which then induce maps

$$\operatorname{coker} f \longrightarrow \operatorname{coker} g \longrightarrow \operatorname{coker} h.$$

To make the notation less heavy, we denote the induced maps on the quotients by i and p. Again, the fact that $pi = 0$ automatically gives us that the restrictions satisfy

$$\operatorname{im}(\operatorname{coker} f \to \operatorname{coker} g) \subseteq \ker(\operatorname{coker} g \to \operatorname{coker} h),$$

so we only need to check equality. Consider $b + \operatorname{im} g$ such that $p(b + \operatorname{im} g) = 0$, meaning that $p(b) = 0$, meaning that $p(b) \in \operatorname{im} h$. Let $c' \in C$ be such that $h(c') = p(b)$. Since p' is surjective, there exists $b' \in B'$ such that $p'(b') = c'$, and by commutativity,

$$pg(b') = hp'(b') = h(c') = p(b).$$

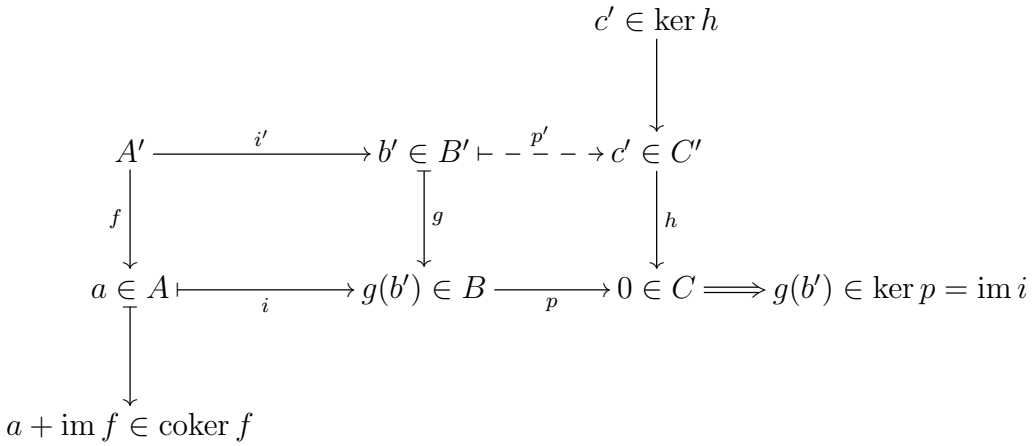
Then $b - g(b') \in \ker p = \operatorname{im} i$. Let $a \in A$ be such that $i(a) = b - g(b')$. Now in $\operatorname{coker} g$ we have

$$\begin{aligned} b + \operatorname{im} g &= b - g(b') + \operatorname{im} g \\ &= i(a) + \operatorname{im} g \\ &= i(a + \operatorname{im} f). \end{aligned}$$

This concludes the proof of exactness of

$$\ker f \longrightarrow \ker g \longrightarrow \ker h \qquad \text{and} \qquad \operatorname{coker} f \longrightarrow \operatorname{coker} g \longrightarrow \operatorname{coker} h.$$

We still need to show the parts of the statement related to the connecting homomorphism ∂. Our definition of ∂ can be visualized as follows:



Let's recap the process in words. First, we fix $c' \in \ker h \subseteq C'$. Since p' is surjective, we can always pick $b' \in B'$ such that $p'(b') = c'$. Since $c' \in \ker h$, by commutativity we have

$$pg(b') = hp'(b') = h(c') = 0,$$

so $g(b') \in \ker p = \operatorname{im} i$. Therefore, there exists $a \in A$ such that $i(a) = g(b')$. In fact, since i is injective, there exists a unique $a \in A$ such that $i(a) = g(b')$. Our definition of $\partial(c')$ sets

$$\partial(c') = a + \operatorname{im} f \in \operatorname{coker} f.$$

The fact that ∂ is a homomorphism of R-modules follows from the fact that all the maps involved are homomorphisms of R-modules: given $c'_1, c'_2 \in \ker h$, and $b'_1, b'_2 \in B'$, $a_1, a_2 \in A$ such that

$$p'(b'_1) = c'_1, \quad p'(b'_2) = c'_2, \quad i(a_1) = g(b'_1), \quad i(a_2) = g(b'_2),$$

we have

$$i(a_1 + a_2) = i(a_1) + i(a_2) = g(b'_1) + g(b'_2) = g(b'_1 + b'_2),$$

so

$$\partial(c'_1) = a_1 + \operatorname{im} f, \quad \partial(c'_2) = a_2 + \operatorname{im} f, \quad \text{and} \quad \partial(c'_1 + c'_2) = (a_1 + a_2) + \operatorname{im} f.$$

Therefore, $\partial(c'_1) + \partial(c'_2) = \partial(c'_1 + c'_2)$. Similarly, given any $r \in R$,

$$r(a_1 + \operatorname{im} f) = ra_1 + \operatorname{im} f, \quad i(ra_1) = ri(a_1) = rg(b'_1) = g(rb'_1), \quad \text{and} \quad p'(rb_1) = rp'(b_1) = rc_1,$$

so $\partial(rc_1) = r(a_1 + \operatorname{im} f) = r\partial(c_1)$. We now need to show the following:

1) ∂ is well-defined. 2) $p'(\ker g) = \ker \partial$. 3) $\operatorname{im} \partial = \ker(\operatorname{coker} f \xrightarrow{i} \operatorname{coker} g)$.

Points 2) and 3) together say that the sequence

$$\ker g \longrightarrow \ker h \xrightarrow{\partial} \operatorname{coker} f \longrightarrow \operatorname{coker} g$$

is exact, and this will complete the proof.

First, let's show that $\partial(0)$ is well-defined. Ultimately, our definition of ∂ only involves one choice, when we pick $b' \in B'$ such that $p'(b') = 0$; we need to show that $\partial(0)$ does not depend on the choice of b'. Given $b' \in B'$ such that $p'(b') = 0$, by exactness we have $b' \in \ker p' = \operatorname{im} i'$. Therefore, there exists $a' \in A'$ such that $i'(a') = b'$. Notice that $a := f(a') \in A$ is such that

$$i(a) = if(a') = gi'(a') = g(b').$$

Thus our definition says that $\partial(0) = a + \operatorname{im} f \in \operatorname{coker} f$. Since $a = f(a') \in \operatorname{im} f$, we conclude that $a + \operatorname{im} f = 0$, so $\partial(0) = 0$ for any choice of b'.

Now consider any $c' \in \ker h$. Again, to show ∂ is well-defined, we need only to show it does not depend on the choice of b' such that $p'(b') = c'$. Consider $b'_1, b'_2 \in B'$ such that

$$p'(b'_1) = p'(b'_2) = c',$$

and $a_1, a_2 \in A$ such that

$$i(a_1) = g(b'_1) \qquad \text{and} \qquad i(a_2) = g(b'_2).$$

Note that

$$i(a_1 - a_2) = g(b'_1 - b'_2),$$

and since

$$p'(b'_1 - b'_2) = c' - c' = 0,$$

we must have

$$a_1 - a_2 + \operatorname{im} f = \partial(0) = 0.$$

Thus

$$a_1 + \operatorname{im} f = a_2 + \operatorname{im} f,$$

and this concludes our proof that ∂ is well-defined.

Now we show 2): that $p'(\ker g) = \ker \partial$.

If $b' \in \ker g$, then the only $a \in A$ such that $i(a) = g(b') = 0$ is $a = 0$. Therefore, $\partial(p'(b')) = 0$, so $p'(\ker g) \subseteq \ker \partial$. On the other hand, let $c' \in \ker h$ be such that $\partial(c') = 0$. That means that for any $b' \in B'$ such that $p'(b') = c'$ we must have $g(b') = i(a)$ for some $a \in \operatorname{im} f$. Let $a' \in A'$ be such that $f(a') = a$. Then

$$gi'(a') = if(a') = i(a) = g(b')$$

so $b' - i'(a') \in \ker g$. Since $p'i' = 0$,

$$c' = p'(b') = p'(b' - i'(a')) \in p'(\ker g).$$

We conclude that $\ker \partial = p'(\ker g)$, and this shows 2).

Now we show 3), that is, $\operatorname{im} \partial = \ker(\operatorname{coker} f \xrightarrow{i} \operatorname{coker} g)$.

Let $a \in A$ be such that $i(a + \operatorname{im} f) = 0$. In B, this says that $i(a) \in \operatorname{im} g$, so we can choose $b' \in B'$ such that $g(b') = i(a)$. Using commutativity and the fact that $pi = 0$, we have

$$hp'(b') = pg(b') = pi(a) = 0 \quad \text{so} \quad p'(b') \in \ker h.$$

This shows that $a + \operatorname{im} f = \partial(p'(b'))$, and thus $\ker(\operatorname{coker} f \xrightarrow{i} \operatorname{coker} g) \subseteq \operatorname{im} \partial$. Finally, if $p'(b') = c'$ and $i(a) = g(b')$, then

$$i\partial(c') = i(a + \operatorname{im} f) = g(b') + \operatorname{im} g = 0, \quad \text{so} \quad \operatorname{im} \partial \subseteq \ker(\operatorname{coker} f \xrightarrow{i} \operatorname{coker} g).$$

Finally, if i' is injective then so is the induced map on $\ker f \longrightarrow \ker g$, since it is obtained by restricting the injective map i'. Similarly, if p is surjective then so is the induced map on $\operatorname{coker} g \longrightarrow \operatorname{coker} h$, as for any $x + \operatorname{im} h$ we can find $y \in B$ such that $p(y) = x$, and so the induced map on the cokernels has

$$y + \operatorname{im} g \longmapsto p(y) + \operatorname{im} h = x + \operatorname{im} h. \qquad \square$$

The proof of the Snake Lemma is what we call a *diagram chase*, for reasons that may be obvious by now: we followed the diagram in the natural way, and everything worked out in the end. The Five Lemma is another classical example of a diagram chase.

Now that we have the Snake Lemma, we can construct the long exact sequence in homology:

Theorem 2.39 (Long exact sequence in homology)**.** *Given a short exact sequence in* $\operatorname{Ch}(R)$

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0,$$

there are connecting homomorphisms $\partial : \mathrm{H}_n(C) \longrightarrow \mathrm{H}_{n-1}(A)$ *such that*

$$\cdots \longrightarrow \mathrm{H}_{n+1}(C) \xrightarrow{\partial} \mathrm{H}_n(A) \xrightarrow{f} \mathrm{H}_n(B) \xrightarrow{g} \mathrm{H}_n(C) \xrightarrow{\partial} \mathrm{H}_{n-1}(A) \longrightarrow \cdots$$

is an exact sequence.

Proof. For each n, we have short exact sequences

$$0 \longrightarrow A_n \longrightarrow B_n \longrightarrow C_n \longrightarrow 0.$$

The condition that f and g are maps of complexes implies, by Theorem 2.5, that f and g take cycles to cycles. This gives us an induced map $Z_n(A) \longrightarrow Z_n(B)$, and since this map is the restriction of an inclusion, it must also be an inclusion. Therefore, we get exact sequences

$$0 \longrightarrow Z_n(A) \longrightarrow Z_n(B) \longrightarrow Z_n(C)\ .$$

Again by Theorem 2.5, the condition that f and g are maps of complexes also implies that f and g both take boundaries to boundaries, so that we get exact sequences

$$A_n/\operatorname{im} d^A_{n+1} \longrightarrow B_n/\operatorname{im} d^B_{n+1} \longrightarrow C_n/\operatorname{im} d^C_{n+1} \longrightarrow 0\ .$$

Let F be any complex. The boundary maps on F induce maps $F_n \longrightarrow Z_{n-1}(F)$ that send $\operatorname{im} d_{n+1}$ to 0, so we get induced maps $F_n/\operatorname{im} d_{n+1} \longrightarrow Z_{n-1}(F)$. Applying this general fact to A, B, and C, and putting all this together, we have a commutative diagram with exact rows

$$\begin{array}{ccccccccc} & & A_n/\operatorname{im} d^A_{n+1} & \longrightarrow & B_n/\operatorname{im} d^B_{n+1} & \longrightarrow & C_n/\operatorname{im} d^C_{n+1} & \longrightarrow & 0\,. \\ & & \downarrow{\scriptstyle d^A_n} & & \downarrow{\scriptstyle d^B_n} & & \downarrow{\scriptstyle d^C_n} & & \\ 0 & \longrightarrow & Z_{n-1}(A) & \longrightarrow & Z_{n-1}(B) & \longrightarrow & Z_{n-1}(C) & & \end{array}$$

For any complex F,

$$\ker(F_n/\operatorname{im} d^F_{n+1} \xrightarrow{d^F_n} Z_{n-1}(F)) = \mathrm{H}_n(F)$$

and

$$\operatorname{coker}(F_n/\operatorname{im} d^F_{n+1} \xrightarrow{d^F_n} Z_{n-1}(F)) = Z_{n-1}(F)/\operatorname{im} d^F_n = \mathrm{H}_{n-1}(F).$$

The Snake Lemma now gives us exact sequences

$$\mathrm{H}_n(A) \longrightarrow \mathrm{H}_n(B) \longrightarrow \mathrm{H}_n(C) \xrightarrow{\partial} \mathrm{H}_{n-1}(A) \longrightarrow \mathrm{H}_{n-1}(B) \longrightarrow \mathrm{H}_{n-1}(C).$$

Finally, we glue all these together to obtain the long exact sequence in homology. □

Remark 2.40. It's helpful to carefully consider how to compute the connecting homomorphisms in the long exact sequence in homology, which we can easily put together from the proof of the Snake Lemma. Suppose that $c \in Z_{n+1}(C) = \ker d^C_{n+1}$. When we view c as an element in C_{n+1}, we can find $b \in B_{n+1}$ such that $g_{n+1}(b) = c$, since g_{n+1} is surjective by assumption. Since g is a map of complexes, we have

$$g_n d^B_{n+1}(b) = d^C_{n+1} g_{n+1}(b) = d^C_{n+1}(c) = 0,$$

so $d^B_{n+1}(b) \in \ker g_n$. In fact, note that $d^B_{n+1}(b) \in Z_n(B)$, so

$$b \in \ker(Z_n(B) \xrightarrow{g_n} Z_n(C)) = \operatorname{im}(Z_n(A) \to Z_n(B)).$$

Thus there exists $a \in Z_n(A)$ such that $f_n(a) = d^B_{n+1}(b)$. Finally,

$$\partial(c + \operatorname{im} d_{n+2}) = a + \operatorname{im} d^A_{n+1}.$$

So in summary, the recipe goes as follows: given $c + \operatorname{im} d_{n+2} \in H_{n+1}(C)$, we find $b \in B_{n+1}$ such that $g_{n+1}(b) = c$ and $a \in Z_n(A)$ such that $f_n(a) = d^B_{n+1}(b)$, and

$$\partial(c) = a + \operatorname{im} d^A_{n+1}.$$

We will soon see that long exact sequences appear everywhere, and that they are very helpful. Before we see more examples, we want to highlight a connection between long and short exact sequences.

Remark 2.41. Suppose that

$$\cdots \longrightarrow C_{n+1} \xrightarrow{f_{n+1}} C_n \xrightarrow{f_n} \cdots$$

is a long exact sequence. This long exact sequence breaks into the short exact sequences

$$0 \longrightarrow \ker f_n \xrightarrow{i} C_n \xrightarrow{\pi} \operatorname{coker} f_{n+1} \longrightarrow 0 \,.$$

The first map i is simply the inclusion of the submodule $\ker f_n$ into C_n, while the second map π is the canonical projection onto the quotient. While it is clear that i is injective and π is surjective, exactness at the middle is less obvious. This follows from the exactness of the original complex, which gives $\operatorname{im} i = \ker f_n = \operatorname{im} f_{n+1} = \ker \pi$.

The long exact sequence in homology is natural.

Theorem 2.42 (Naturality of the long exact sequence in homology)**.** *Any commutative diagram in* $\mathrm{Ch}(R)$

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{i} & B & \xrightarrow{p} & C & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle h} & & \\ 0 & \longrightarrow & A' & \xrightarrow[i']{} & B' & \xrightarrow[p']{} & C' & \longrightarrow & 0 \end{array}$$

with exact rows induces a commutative diagram with exact rows

$$\begin{array}{ccccccccccccc} \cdots & \longrightarrow & \mathrm{H}_{n+1}(C) & \xrightarrow{\partial} & \mathrm{H}_n(A) & \xrightarrow{i} & \mathrm{H}_n(B) & \xrightarrow{p} & \mathrm{H}_n(C) & \xrightarrow{\partial} & \mathrm{H}_{n-1}(A) & \longrightarrow & \cdots \\ & & \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle f} & & \\ \cdots & \longrightarrow & \mathrm{H}_{n+1}(C') & \xrightarrow[\partial']{} & \mathrm{H}_n(A') & \xrightarrow[i']{} & \mathrm{H}_n(B') & \xrightarrow[p']{} & \mathrm{H}_n(C') & \xrightarrow[\partial']{} & \mathrm{H}_{n-1}(A') & \longrightarrow & \cdots \end{array}$$

Proof. The rows of the resulting diagram are the long exact sequences in homology induced by each row of the original diagram, as in Theorem 2.39. So the content of the theorem is that the maps induced in homology by f, g, and h make the diagram commute. The commutativity of

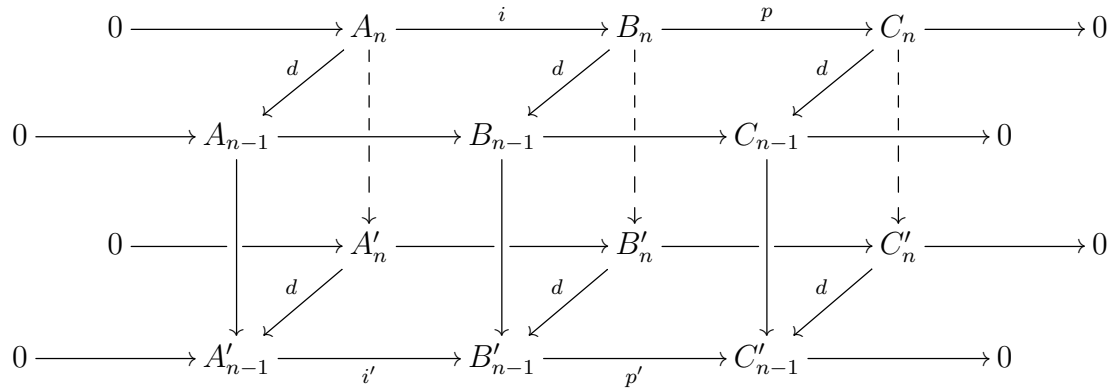
$$\begin{array}{ccccc} \mathrm{H}_n(A) & \xrightarrow{i} & \mathrm{H}_n(B) & \xrightarrow{p} & \mathrm{H}_n(C) \\ \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle h} \\ \mathrm{H}_n(A') & \xrightarrow[i']{} & \mathrm{H}_n(B') & \xrightarrow[p']{} & \mathrm{H}_n(C') \end{array}$$

follows from the fact that H_n is a functor, so we only need to check commutativity of the square

$$\begin{array}{ccc} \mathrm{H}_n(C) & \xrightarrow{\partial} & \mathrm{H}_{n-1}(A) \\ \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle f} \\ \mathrm{H}_n(C') & \xrightarrow[\partial']{} & \mathrm{H}_{n-1}(A') \end{array}$$

that involves the connecting homomorphisms ∂ and ∂'.

Consider the following commutative diagram:



Given $c \in \ker(d_n \colon C_n \longrightarrow C_{n-1})$, we need to check that $f_{n-1}(\partial(c)) = \partial' h_n(c)$ in $\mathrm{H}_{n-1}(A')$. To compute $\partial(c)$, we find a lift $b \in B_n$ such that $p_n(b) = c$, and $a \in A_{n-1}$ with $i_{n-1}(a) = d_n(b)$, and set $\partial(c) = a + \operatorname{im} d_n \in \mathrm{H}_{n-1}(A)$. So $f_{n-1}\partial(c) = f_{n-1}(a) + \operatorname{im} d_n$. On the other hand, to compute $\partial' h_n(c)$, we start by finding $b' \in B'_n$ such that $p'_n(b') = h_n(c)$. By commutativity of the right back square

$$\begin{array}{ccc} B_n & \xrightarrow{p_n} & C_n \\ {\scriptstyle g_n}\downarrow & & \downarrow{\scriptstyle h_n} \\ B'_n & \xrightarrow[p'_n]{} & C'_n \end{array}$$

we can choose $b' = g_n(b)$, since

$$p'_n(b') = p'_n g_n(b) = h_n p_n(b) = h_n(c).$$

Next we take $a' \in A'_{n-1}$ such that $i'_{n-1}(a') = d_n(b')$, and set $\partial'(h(c)) = a' + \operatorname{im} d_n \in \mathrm{H}_{n-1}(A')$. By commutativity of the middle square

$$\begin{array}{ccc} B_n & \xrightarrow{d_n} & B_{n-1} \\ {\scriptstyle g_n}\downarrow & & \downarrow{\scriptstyle g_{n-1}} \\ B'_n & \xrightarrow[d_n]{} & B'_{n-1} \end{array}$$

we have

$$d_n(b') = d_n g_n(b) = g_{n-1} d_n(b).$$

By our choice of a, we have

$$d_n(b') = g_{n-1} d_n(b) = g_{n-1} i_{n-1}(a),$$

and by commutativity of the front left square

$$\begin{array}{ccc} A_{n-1} & \xrightarrow{i_{n-1}} & B_{n-1} \\ {\scriptstyle f_{n-1}}\downarrow & & \downarrow{\scriptstyle g_{n-1}} \\ A'_{n-1} & \xrightarrow[i'_{n-1}]{} & B'_{n-1} \end{array} \qquad \text{we have} \quad i'_{n-1} f_{n-1}(a) = g_{n-1} i_{n-1}(a) = d_n(b').$$

So we can take $a' = f_{n-1}(a)$. Finally, this means $\partial'(h_n(c)) = f_{n-1}(a) + \operatorname{im} d_{n-1}$, as desired. $\square$

Remark 2.43. Let

$$0 \longrightarrow A \xrightarrow{i} B \xrightarrow{p} C \longrightarrow 0$$

be a short exact sequence in $\mathrm{Ch}(R)$. We can think of Theorem 2.42 as saying that the induced maps on homology $i_*\colon \mathrm{H}_n(A) \longrightarrow \mathrm{H}_n(B)$ and $p_*\colon \mathrm{H}_n(B) \longrightarrow \mathrm{H}_n(C)$ and the connecting homomorphism $\partial\colon \mathrm{H}_n(C) \longrightarrow \mathrm{H}_{n-1}(A)$ are all natural transformations. More precisely, consider the category **SES** of short exact sequences of R-modules, which is a full subcategory of $\mathrm{Ch}(R)$. Homology gives us functors **SES** $\longrightarrow$ R-**Mod** that given a short exact sequence

$$0 \longrightarrow A \xrightarrow{i} B \xrightarrow{p} C \longrightarrow 0$$

return the R-modules $\mathrm{H}_n(A)$, $\mathrm{H}_n(B)$, or $\mathrm{H}_n(C)$). A map between two short exact sequences then induces R-module homomorphisms between the corresponding homologies. With this framework, Theorem 2.42 says that $i_*\colon \mathrm{H}_n(A) \longrightarrow \mathrm{H}_n(B)$, and $p_*\colon \mathrm{H}_n(B) \longrightarrow \mathrm{H}_n(C)$ and the connecting homomorphism $\partial : \mathrm{H}_n(C) \longrightarrow \mathrm{H}_{n-1}(A)$ are all natural transformations between the corresponding homology functors.

Chapter 3

R-Mod

Before we study abelian categories in general, we want to understand our best prototype for what an abelian category looks like: the category R-**Mod** of R-modules and R-module homomorphisms.

3.1 Hom

From now on, let's fix a ring R. Recall that whenever we say an R-module M, we mean a *left* R-module; any general facts about left modules can be naturally converted into statements about right R-modules, under small appropriate corrections. When M is commutative, left and right module structures agree, so the distinction is not relevant.

Our goal is to get to know the category R-**Mod**, which as we are about to discover is a very nice category. One of the many nice things about R-**Mod** is that the Hom-sets have an extra structure. (Roughly speaking, a locally small category where the Hom-sets are objects in some other category is called an *enriched category*).

To make the notation less heavy, we write $\mathrm{Hom}_R(M, N)$ instead of $\mathrm{Hom}_{R\text{-}\mathbf{Mod}}(M, N)$ for the Hom-set between M and N in R-**Mod**. The arrows in $\mathrm{Hom}_R(M, N)$ are all the R-module homomorphisms from M to N. This is a locally small category, meaning that the Hom-sets are actual sets, but more even is true: the Hom-sets are actually abelian groups, and when R is commutative, they are even R-modules.

Given $f, g \in \mathrm{Hom}_R(M, N)$, $f + g$ is the R-module homomorphism defined by

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

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

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

Exercise 3.1. Let M and N be R-modules. Then $\mathrm{Hom}_R(M, N)$ is an abelian group under the sum defined above.

Exercise 3.2. Let M and N be R-modules over a commutative ring R. Then $\mathrm{Hom}_R(M, N)$ is an R-module.

Remark 3.3. The main reason we need commutativity for $\mathrm{Hom}_R(M,N)$ to be a module is that given any $r \in R$ and $f \in \mathrm{Hom}_R(M,N)$, we need rf to be an R-module homomorphism, so in particular for any $a \in M$ and any $s \in R$ we need

$$(rf)(sa) = s(rf)(a),$$

so

$$(rs)f(a) = rf(sa) = (rf)(sa) = s(rf)(a) = s(rf(a)) = (sr)f(a).$$

This holds whenever $rs = sr$, but not in general.

Some Hom-sets can easily be identified with other well-understood modules.

Exercise 3.4. Let R be a commutative ring. Let M be an R-module, and I an ideal in R. Then we have the following isomorphisms of R-modules:

a) $\mathrm{Hom}_R(R,M) \cong M$.

b) $\mathrm{Hom}_R(R^n,M) \cong M^n$ for any $n \geqslant 1$.

c) $\mathrm{Hom}_R(R/I,M) \cong (0 :_M I) := \{m \in M \mid Im = 0\}$.

In fact, part a) above can be upgraded to the following:

Exercise 3.5. Let R be a commutative ring. Then $\mathrm{Hom}_R(R,-)$ is naturally isomorphic to the identity functor on R-**Mod**.

Since R-**Mod** is a locally small category, we saw in Theorem 1.43 that there are two Hom-functors from R-**Mod** to **Set**, the covariant functor $\mathrm{Hom}_R(M,-)\colon R\text{-}\mathbf{Mod} \longrightarrow \mathbf{Set}$ and the contravariant functor $\mathrm{Hom}_R(-,N)\colon R\text{-}\mathbf{Mod} \longrightarrow \mathbf{Set}$. In light of Theorem 3.2, we can upgrade these functors to land in **Ab**, or in R-**Mod** when R is commutative, not just in **Set**. Note that while there are two Hom-functors, we will sometimes refer to *the* Hom functor when talking about properties that are common to both of them.

A functor that lands in R-**Mod**, or **Ab** in particular, can have some additional good properties.

Definition 3.6. Let R and S be rings. A functor $T\colon R\text{-}\mathbf{Mod} \longrightarrow S\text{-}\mathbf{Mod}$ is an **additive functor** if

$$T(f+g) = T(f) + T(g)$$

for all $f, g \in \mathrm{Hom}_R(M,N)$.

To say that T is a covariant additive functor is to say that for all A and B, the map

$$\begin{aligned} \mathrm{Hom}(A,B) &\longrightarrow \mathrm{Hom}(T(A),T(B)) \\ f &\longmapsto T(f) \end{aligned}$$

induced by T is a homomorphism of abelian groups. Similarly, a contravariant additive functor T is one such that

$$\begin{aligned} \mathrm{Hom}(A,B) &\longrightarrow \mathrm{Hom}(T(B),T(A)) \\ f &\longmapsto T(f) \end{aligned}$$

is a homomorphism of abelian groups. Notice moreover that this definition makes sense more generally in any category $\mathscr{C}$ whose objects have an abelian group structure.

Exercise 3.7. Show that $\mathrm{Hom}_R(M, -)$ and $\mathrm{Hom}_R(-, N)$ are both additive functors.

Note that in the previous exercise we were purposely vague about where $\mathrm{Hom}_R(M, -)$ and $\mathrm{Hom}_R(-, N)$ land: these are additive functors whether we consider them as functors with target **Ab** or target R-**Mod**, when appropriate.

Additive functors have many nice properties.

Lemma 3.8. *Let $T\colon R$-**Mod** $\longrightarrow S$-**Mod** be an additive functor.*

a) Let 0 denote the 0-map between any two R-modules M and N. Then $T(0) = 0$ is the 0-map $T(M) \to T(N)$.

b) Let 0 denote the zero R-module. Then $T(0) = 0$ is the zero S-module.

Proof.

a) As a function defined on each fixed $\mathrm{Hom}_R(M, N)$, T is a group homomorphism, so it must send 0 to 0.

b) An R-module M is the zero module if and only if the zero and identity maps on M coincide. Let N be the image of the zero R-module via T. On the one hand, any functor must send identity maps to identity maps, so the identity map on the zero module must be sent to the identity on N. On the other hand, we have shown that the zero map must be sent to the zero map on N, so the zero and identity maps on N must coincide, so $N = 0$. □

Remark 3.9. Note that the category of chain complexes also has a similar structure to R-**Mod**: given two maps of complexes $f, g\colon C \to D$, we define a map of complexes $f + g\colon C \to D$ given by

$$(f + g)_n := f_n + g_n.$$

It is routine to check that this again gives a map of complexes, and that this operation gives the Hom-sets in $\mathrm{Ch}(R)$ the structure of an abelian group. In fact, this abelian group structure can be upgraded to an R-module structure when R is commutative, by setting

$$(rf)_n := rf_n$$

for all $r \in R$. This allows us to talk about additive functors to and from the category $\mathrm{Ch}(R)$, and there is a version of Theorem 3.8 in $\mathrm{Ch}(R)$.

Exercise 3.10. Show that homology is an additive functor.

Most functors between categories or modules or chain complexes are additive. In fact, we will spend the rest of this chapter studying three very important additive functors: the two Hom functors, and a third functor we have yet to define.

Exercise 3.11. Let R and S be rings and let $T : R$-**Mod** $\longrightarrow S$-**Mod** be an additive functor. Show that for all R-modules A and B,

$$T(A \oplus B) \cong T(A) \oplus T(B).$$

Hom satisfies a stronger version of this property.

Theorem 3.12. *For all R-modules M, N, M_i, N_i, there are isomorphisms of abelian groups*

$$\mathrm{Hom}_R(M, \prod_i N_i) \cong \prod_i \mathrm{Hom}_R(M, N_i) \text{ and } \mathrm{Hom}_R(\bigoplus_i M_i, N) \cong \prod_i \mathrm{Hom}_R(M_i, N).$$

Moreover, when R is commutative, these are in fact isomorphisms of R-modules.

In particular,

$$\mathrm{Hom}_R(A \oplus B, C) \cong \mathrm{Hom}_R(A, C) \oplus \mathrm{Hom}_R(B, C)$$

and

$$\mathrm{Hom}_R(A, B \oplus C) \cong \mathrm{Hom}_R(A, B) \oplus \mathrm{Hom}_R(A, C).$$

These two properties, however, are consequences of Theorem 3.7 and Theorem 3.11: Hom is additive, and additive functors preserve finite direct sums.

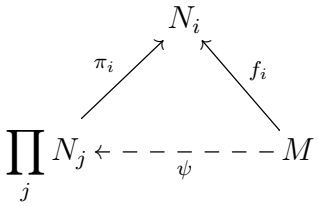
Proof. For each i, let $\pi_i : \prod_j N_j \longrightarrow N_i$ be the canonical projection map. Consider the map

$$\begin{array}{ccc} \mathrm{Hom}_R(M, \prod_i N_i) & \xrightarrow{\alpha} & \prod_i \mathrm{Hom}_R(M, N_i) \\ f & \longmapsto & (\pi_i f) \end{array}.$$

We claim this map is the desired isomorphism. We leave it as an exercise to show that α is a homomorphism of abelian groups, and a homomorphism of R-modules when R is commutative; we focus on proving that α is a bijection. First, take $(f_i)_i \in \prod_i \mathrm{Hom}_R(M, N_i)$. Define a map

$$\begin{array}{ccc} M & \xrightarrow{\psi} & \prod_i N_i \\ m & \longmapsto & (f_i(m)) \end{array}.$$

This makes the diagram



commute, so that $\alpha(\psi) = (\pi_i \psi)_i = (f_i)$. This shows that α us surjective.

Now let us show that α is injective. Suppose $f \in \mathrm{Hom}_R(M, \prod_i N_i)$ is such that $\alpha(f) = 0$. For each $m \in M$, let $f(m) = (n_i)_i$, so $\pi_i f(m) = n_i$. By assumption, $(\pi_i f(m)) = 0$, which means that $\pi_i \alpha = 0$ for all i, and thus $n_i = 0$ for all i. So $f = 0$. We conclude that α is an isomorphism.

Now consider the map

$$\begin{array}{ccc} \mathrm{Hom}_R(\bigoplus_i M_i, N) & \xrightarrow{\beta} & \prod_i \mathrm{Hom}_R(M_i, N) \\ f & \longmapsto & (f \iota_i) \end{array}$$

where $\iota_j \colon M_j \longrightarrow \bigoplus_i M_i$ is the inclusion of the jth factor. We leave it as an exercise to prove that β is a homomorphism of abelian groups, and that whenever R is commutative, β is in fact a homomorphism of R-modules.

Given $(f_i)_i \in \prod_i \mathrm{Hom}_R(M_i, N)$, let

$$\begin{aligned} \bigoplus_i M_i &\xrightarrow{\quad\psi\quad} N \\ (m_i) &\longmapsto \sum_i f_i(m_i) \end{aligned}$$

Then $\beta(\psi) = (\psi\iota_i)_i$, so for each i and each $m_i \in M_i$, $\psi\iota_i(m_i) = f_i(m_i)$, and $\beta(\psi) = (f_i)_i$. This shows that β is surjective.

Now assume $\beta(f) = 0$, which implies that $f\iota_i$ is the zero map for each i. Consider any $(m_i)_i \in \bigoplus_i M_i$. For each i, $f\iota_i(m_i) = 0$. On the other hand, $(m_i)_i = \sum_i \iota_i(m_i)$, so $f((m_i)_i) = \sum_i \iota_i(m_i) = 0$. We conclude that $f = 0$, and β is injective. □

Exercise 3.13. Show that the isomorphisms in Theorem 3.12 are natural on both components. More precisely, given any other family of R-modules L_i such that for each i there exists j, a map σ_{ij} there exist R-module maps making the following diagrams commute:

$$\begin{array}{ccc} \mathrm{Hom}_R(M, \prod_i N_i) & \xrightarrow{\cong} & \prod_i \mathrm{Hom}_R(M, N_i) \\ \downarrow & & \downarrow \\ \mathrm{Hom}_R(M, \prod_i L_i) & \xrightarrow[\cong]{} & \prod_i \mathrm{Hom}_R(M, L_i) \end{array} \qquad \begin{array}{ccc} \mathrm{Hom}_R(\bigoplus_i M_i, N) & \xrightarrow{\cong} & \bigoplus_i \mathrm{Hom}_R(M_i, N) \\ \downarrow & & \downarrow \\ \mathrm{Hom}_R(\bigoplus_i L_i, N) & \xrightarrow[\cong]{} & \bigoplus_i \mathrm{Hom}_R(L_i, N) \end{array}$$

$$\begin{array}{ccc} \mathrm{Hom}_R(M, \prod_i N_i) & \xrightarrow{\cong} & \prod_i \mathrm{Hom}_R(M, N_i) \\ \downarrow & & \downarrow \\ \mathrm{Hom}_R(M, \prod_i L_i) & \xrightarrow[\cong]{} & \prod_i \mathrm{Hom}_R(M, L_i) \end{array} \qquad \begin{array}{ccc} \mathrm{Hom}_R(\bigoplus_i M_i, N) & \xrightarrow{\cong} & \bigoplus_i \mathrm{Hom}_R(M_i, N) \\ \downarrow & & \downarrow \\ \mathrm{Hom}_R(\bigoplus_i L_i, N) & \xrightarrow[\cong]{} & \bigoplus_i \mathrm{Hom}_R(L_i, N) \end{array}$$

In fact, one can show that more generally, Hom behaves well with limits and colimits.

Exercise 3.14. Let R be any ring and consider R-modules A and $\{M_i\}$.

a) For any inverse system $\{M_i\}$, there is a natural isomorphism

$$\mathrm{Hom}_R(A, \varprojlim_i M_i) \cong \varprojlim_i \mathrm{Hom}_R(A, M_i).$$

b) For any direct system $\{M_i\}$ or R-modules, there is a natural isomorphism

$$\mathrm{Hom}_R(\mathrm{colim}_i\, M_i, A) \cong \varprojlim_i \mathrm{Hom}_R(M_i, B).$$

Moreover, when R is commutative, these are isomorphisms of modules.

Another important property of Hom is how it interacts with exact sequences. First, an important note about general additive functors:

Remark 3.15. Let $F\colon R\textbf{-Mod} \to S\textbf{-Mod}$ be an additive functor. Thanks to Theorem 3.8, if $gf = 0$, then

$$F(gf) = F(g)F(f) = F(0) = 0.$$

Thus F must send complexes to complexes, and in fact, F induces a functor $\mathrm{Ch}(R) \to \mathrm{Ch}(S)$, which we also call F. Now if h is a homotopy between two maps of complexes, F must preserve the identities

$$\delta_{n+1}h_n + h_{n-1}\delta_n = f_n - g_n$$

for all n, so $F(h)$ is a homotopy between $F(f)$ and $F(g)$.

While additive functors send complexes to complexes, they don't have to preserve exactness. Functors that do preserve exactness are very special.

Definition 3.16. An additive functor $T : R\textbf{-Mod} \longrightarrow S\textbf{-Mod}$ is an **exact functor** if it preserves short exact sequences. When T is covariant, this means that every short exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

is taken to the short exact sequence

$$0 \longrightarrow T(A) \xrightarrow{T(f)} T(B) \xrightarrow{T(g)} T(C) \longrightarrow 0.$$

When T is contravariant, this means that any short exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

is taken to the short exact sequence

$$0 \longrightarrow T(C) \xrightarrow{T(g)} T(B) \xrightarrow{T(f)} T(A) \longrightarrow 0.$$

Exercise 3.17. Show that an additive functor T is exact if and only if it commutes with homology, that is, for all complexes C and all n,

$$\mathrm{H}_n(T(C)) = T(\mathrm{H}_n(C)).$$

As we will soon see, most functors are not exact. However, many functors of interest preserve some exactness.

Definition 3.18. A covariant additive functor $T : R\textbf{-Mod} \longrightarrow S\textbf{-Mod}$ is **left exact** if it takes every exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C$$

of R-modules to the exact sequence

$$0 \longrightarrow T(A) \xrightarrow{T(f)} T(B) \xrightarrow{T(g)} T(C)$$

of S-modules, and **right exact** if it takes every exact sequence of R-modules

$$A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

to the exact sequence of S-modules

$$T(A) \xrightarrow{T(f)} T(B) \xrightarrow{T(g)} T(C) \longrightarrow 0.$$

Definition 3.19. A contravariant additive functor $T\colon R\text{-}\mathbf{Mod} \longrightarrow S\text{-}\mathbf{Mod}$ is **left exact** if it takes every exact sequence

$$A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

of R-modules to the exact sequence

$$0 \longrightarrow T(C) \xrightarrow{T(g)} T(B) \xrightarrow{T(f)} T(A)$$

of S-modules, and **right exact** if it takes every exact sequence of R-modules

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C$$

to the exact sequence of S-modules

$$T(C) \xrightarrow{T(g)} T(B) \xrightarrow{T(f)} T(A) \longrightarrow 0.$$

Exercise 3.20. The definitions above all stay unchanged if for each condition we start with a short exact sequence. For example, a covariant additive functor T is left exact if and only if for every short exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

of R-modules,

$$0 \longrightarrow T(A) \xrightarrow{T(f)} T(B) \xrightarrow{T(g)} T(C)$$

is exact.

Remark 3.21. Left exact covariant functors take kernels to kernels, while right exact covariant functors take cokernels to cokernels: the kernel of f fits in an exact sequence

$$0 \longrightarrow \ker f \longrightarrow A \xrightarrow{f} B$$

and applying a left exact functor F gives us an exact sequence

$$0 \longrightarrow F(\ker f) \longrightarrow F(A) \xrightarrow{F(f)} F(B).$$

Exactness tells us that $F(\ker f)$ is the kernel of $F(f)$. Similarly, the cokernel of f fits into an exact sequence

$$A \xrightarrow{f} B \longrightarrow \operatorname{coker} f \longrightarrow 0,$$

which any right exact functor G will take to an exact sequence

$$G(A) \xrightarrow{G(f)} G(B) \longrightarrow G(\operatorname{coker} f) \longrightarrow 0.$$

Exactness says that $G(\operatorname{coker} f)$ is the cokernel of $G(f)$.

Similarly, left exact contravariant functors take cokernels to kernels, and right exact contravariant functors take kernels to cokernels. A left exact contravariant functor F will take the exact sequence

$$A \xrightarrow{f} B \longrightarrow \operatorname{coker} f \longrightarrow 0$$

to an exact sequence

$$0 \longrightarrow F(\operatorname{coker} f) \longrightarrow F(B) \xrightarrow{F(f)} F(A),$$

and exactness tells us that $F(\operatorname{coker} f)$ is the kernel of $F(f)$.

A right exact contravariant functor G will take the exact sequence

$$0 \longrightarrow \ker f \longrightarrow A \xrightarrow{f} B$$

to the exact sequence

$$G(B) \xrightarrow{G(f)} G(A) \longrightarrow G(\ker f) \longrightarrow 0,$$

and exactness says that $G(\ker f)$ is the cokernel of $G(f)$.

Exactness is preserved by natural isomorphisms.

Remark 3.22. Suppose that $F, G\colon R\textbf{-Mod} \longrightarrow S\textbf{-Mod}$ are naturally isomorphic additive functors. We claim that F is exact if and only if G is exact. Let's prove it in the case when F and G are covariant. Given any short exact sequence

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

applying each of our functors yields complexes of R-modules which may or may not be exact. Our natural isomorphism gives us an isomorphism of complexes

$$\begin{array}{ccccccccc} 0 & \longrightarrow & F(A) & \longrightarrow & F(B) & \longrightarrow & F(C) & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & G(A) & \longrightarrow & G(B) & \longrightarrow & G(C) & \longrightarrow & 0. \end{array}$$

Isomorphisms of complexes induce isomorphisms in homology, so the top sequence is exact if and only if the bottom sequence is exact. Thus F preserves the short exact sequence if and only if G does.

A similar argument shows that F is left (respectively, right) exact if and only if G is left (respectively, right) exact; we leave the details as an exercise.

However, an additive functor does not have to be left exact nor right exact. There are even some functors that preserve exactness *in the middle*.

Example 3.23. The homology functor is exact *in the middle*: given a short exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0,$$

the exactness of the long exact sequence in homology says in particular that

$$\mathrm{H}_n(A) \xrightarrow{\mathrm{H}_n(f)} \mathrm{H}_n(B) \xrightarrow{\mathrm{H}_n(g)} \mathrm{H}_n(C)$$

is exact for all n. On the other hand, we claim that the homology functor is neither left exact nor right exact. More precisely, $\mathrm{H}_n(f)$ might fail to be injective and $\mathrm{H}_n(g)$ might fail to be surjective. Finding a counterexample amounts to finding a short exact sequence of complexes such that the connecting homomorphism in the long exact sequence in homology is not the zero map.

For example, consider the following complexes and maps of complexes:

$$\begin{array}{cccccccccc}
 & & & & 2 & & 1 & & 0 & & -1 & \\
A = & \cdots & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \mathbb{Z} & \longrightarrow & 0 & \longrightarrow & \cdots \\
f\downarrow & & & \downarrow & & \downarrow & & \| & & \downarrow & & \\
B = & \cdots & \longrightarrow & 0 & \longrightarrow & \mathbb{Z} & = & \mathbb{Z} & \longrightarrow & 0 & \longrightarrow & \cdots \\
g\downarrow & & & \downarrow & & \| & & \downarrow & & \downarrow & & \\
C = & \cdots & \longrightarrow & 0 & \longrightarrow & \mathbb{Z} & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \cdots .
\end{array}$$

Applying H_0 gives us

$$\begin{array}{c}
\mathrm{H}_0(A) \xrightarrow{\mathrm{H}_0(f)} \mathrm{H}_0(B) \\
\mathbb{Z} \xrightarrow{0} 0,
\end{array}$$

which is not injective, so

$$0 \longrightarrow \mathrm{H}_0(A) \xrightarrow{\mathrm{H}_0(f)} \mathrm{H}_0(B) \xrightarrow{\mathrm{H}_0(g)} \mathrm{H}_0(C)$$

is not exact. Similarly, applying H_1 gives

$$\begin{array}{c}
\mathrm{H}_1(B) \xrightarrow{\mathrm{H}_1(g)} \mathrm{H}_1(C) \\
0 \xrightarrow{0} \mathbb{Z},
\end{array}$$

which is not surjective, so

$$\mathrm{H}_1(A) \xrightarrow{\mathrm{H}_1(f)} \mathrm{H}_1(B) \xrightarrow{\mathrm{H}_1(g)} \mathrm{H}_1(C) \longrightarrow 0$$

is not exact. Thus homology is neither left exact nor right exact, though it is exact in the middle.

But in general, an additive functor might fail to preserve exactness even *in the middle*.

Example 3.24. Fix a prime p and consider the functor $F\colon \mathbf{Ab} \to \mathbf{Ab}$ which on objects is defined by

$$F(M) = \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/p, M/p^2M);$$

given a homomorphism of abelian groups $f\colon M \to N$, we get an induced homomorphism of abelian groups

$$\begin{aligned} M/p^2M &\xrightarrow{\overline{f}} N/p^2N \\ m + p^2M &\longmapsto f(m) + p^2N, \end{aligned}$$

and $F(f) = \overline{f} \circ -$ is postcomposition with $\overline{f}$. Consider the short exact sequence

$$0 \longrightarrow \mathbb{Z}/p^2 \xrightarrow{f} \mathbb{Z}/p^3 \xrightarrow{g} \mathbb{Z}/p \longrightarrow 0,$$

where f is the multiplication by p map, which sends $1 \mapsto p$, and g is the canonical quotient map by the subgroup generated by p.

Note that

$$F(\mathbb{Z}/p^2) = \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/p, \mathbb{Z}/p^2)$$

is the submodule of $\mathbb{Z}/p^2$ of elements killed by p, which is generated by the class of p, so $F(\mathbb{Z}/p^2) = \mathbb{Z}/p$. Moreover,

$$\frac{\mathbb{Z}/p^3}{p^2\mathbb{Z}/p^3} \cong \mathbb{Z}/p^2,$$

so $F(\mathbb{Z}/p^3)$ is the the submodule of $\mathbb{Z}/p^2$ of elements killed by p, which is generated by p and isomorphic to $\mathbb{Z}/p$, so $F(\mathbb{Z}/p^3) = \mathbb{Z}/p$. Now

$$F(f)\colon \mathbb{Z}/p \to \mathbb{Z}/p$$

is the map induced by multiplication by p, so it is the zero map. The map

$$\overline{g}\colon \mathbb{Z}/p^2 \to \mathbb{Z}/p$$

is the canonical quotient by the subgroup generated by p; any element in

$$F(Z/p^3) = \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/p, \mathbb{Z}/p^2)$$

corresponds to choosing an element of order p, and thus in the subgroup generated by p, so applying $\overline{g}$ always results in 0. We conclude that $F(g) = 0$. Finally, this shows that applying F to the original short exact sequence gives us the complex

$$0 \longrightarrow \mathbb{Z}/p \xrightarrow{0} \mathbb{Z}/p \xrightarrow{0} \mathbb{Z}/p \longrightarrow 0,$$

which is not exact anywhere.

One amazing fact, however, is that even if a functor is not exact, it must always preserve *split* short exact sequences.

Exercise 3.25. Show that additive functors preserve split short exact sequences.

We are now ready for our first important example of a left exact functor: Hom is left exact.

Theorem 3.26. *Let M be an R-module.*

a) The covariant functor $\mathrm{Hom}_R(M,-)$ is left exact: for every exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C$$

of R-modules, the sequence

$$0 \longrightarrow \mathrm{Hom}_R(M,A) \xrightarrow{\mathrm{Hom}_R(M,f)} \mathrm{Hom}_R(M,B) \xrightarrow{\mathrm{Hom}_R(M,g)} \mathrm{Hom}_R(M,C)$$

is exact.

b) The contravariant functor $\mathrm{Hom}_R(-,M)$ is left exact: for every exact sequence

$$A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

of R-modules, the sequence

$$0 \longrightarrow \mathrm{Hom}_R(C,M) \xrightarrow{\mathrm{Hom}_R(g,M)} \mathrm{Hom}_R(B,M) \xrightarrow{\mathrm{Hom}_R(f,M)} \mathrm{Hom}_R(A,M)$$

is exact.

Proof. To make the notation less heavy, we will write

$$f_* := \mathrm{Hom}_R(M,f) \quad \text{and} \quad g_* := \mathrm{Hom}_R(M,g)$$

and similarly

$$f^* := \mathrm{Hom}_R(f,M) \quad \text{and} \quad g^* := \mathrm{Hom}_R(g,M).$$

Since additive functors send complexes to complexes, as outlined in Theorem 7.47, we at least know that

$$0 \longrightarrow \mathrm{Hom}_R(M,A) \xrightarrow{\mathrm{Hom}_R(M,f)} \mathrm{Hom}_R(M,B) \xrightarrow{\mathrm{Hom}_R(M,g)} \mathrm{Hom}_R(M,C)$$

and

$$0 \longrightarrow \mathrm{Hom}_R(C,M) \xrightarrow{\mathrm{Hom}_R(g,M)} \mathrm{Hom}_R(B,M) \xrightarrow{\mathrm{Hom}_R(f,M)} \mathrm{Hom}_R(A,M)$$

are functors, so in particular

$$g_* f_* = 0 \implies \mathrm{im}\, f_* \subseteq \ker g_*$$

and

$$f^* g^* = 0 \implies \mathrm{im}\, g^* \subseteq \ker f^*.$$

a) We have two things to show:

f_* is injective:

Suppose that $h \in \mathrm{Hom}_R(M, A)$ is such that $f_*(h) = 0$. By definition, this means that $fh = 0$. But f is injective, so for any $m \in M$

$$fh(m) = 0 \implies h(m) = 0.$$

We conclude that $h = 0$, and f_* is injective.

$\ker g_* \subseteq \mathrm{im}\, f_*$:

Let $h \in \mathrm{Hom}_R(M, B)$ be in the kernel of g_*. Then $gh = g_*(h) = 0$, so for each $m \in M$, $gh(m) = 0$. Then $h(m) \in \ker g = \mathrm{im}\, f$, so there exists $a \in A$ such that $f(a) = h(m)$. Since f is injective, this element a is unique for each $m \in M$. So setting $k(m) := a$ gives us a well-defined function $k \colon M \longrightarrow A$. We claim that k is in fact an R-module homomorphism. To see that, notice that if $k(m_1) = a_1$ and $k(m_2) = a_2$, then

$$f(a_1 + a_2) = f(a_1) + f(a_2) = h(m_1) + h(m_2) = h(m_1 + m_2),$$

so that $k(m_1 + m_2) = a_1 + a_2 = k(m_1) + k(m_2)$. Similarly, given any $r \in R$,

$$f(ra_1) = rf(a_1) = rh(m_1) = h(rm_1),$$

so $k(rm_1) = ra_1 = rk(m_1)$. Finally, this element $k \in \mathrm{Hom}_R(M, A)$ satisfies

$$f_*(k)(m) = f(k(m)) = h(m)$$

for all $m \in M$, so $f_*(k) = h$ and $h \in \mathrm{im}\, f_*$.

b) Again, we have two things to show:

g^* is injective:

If $g^*(h) = 0$ for some $h \in \mathrm{Hom}_R(C, M)$, then $hg = g^*(h) = 0$. Consider any $c \in C$. Since g is surjective, there exists $b \in B$ such that $g(b) = c$. Then $h(c) = hg(b) = 0$, so $h = 0$.

$\ker f^* \subseteq \mathrm{im}\, g^*$:

Let $h \in \mathrm{Hom}_R(B, M)$ be in $\ker f^*$, so that $hf = 0$. Given any $c \in C$, there exists $b \in B$ such that $g(b) = c$, since g is surjective. Let $k \colon C \longrightarrow M$ be the function defined by $k(c) := h(b)$ for some b with $g(b) = c$. This function is well-defined, since whenever $g(b') = g(b) = c$, $b - b' \in \ker g = \mathrm{im}\, f$, say $b - b' = f(a)$, and thus $h(b - b') = h(f(a)) = 0$. Moreover, we claim that k is indeed a homomorphism of R-modules. If $c_1, c_2 \in C$, and $g(b_1) = c_1$, $g(b_2) = c_2$, then $g(b_1 + b_2) = c_1 + c_2$, so

$$k(c_1 + c_2) = h(b_1 + b_2) = h(b_1) + h(b_2) = k(b_1) + k(b_2).$$

Finally, this element $k \in \mathrm{Hom}_R(C, M)$ is such that $g^*(k)$ satisfies

$$(g_*(k))(b) = k(g(b)) = h(b)$$

for all $b \in B$, so $g^*(k) = h$, and $h \in \mathrm{im}\, g^*$. $\square$

So $\mathrm{Hom}_R(M,-)$ preserves kernels, and $\mathrm{Hom}_R(-,N)$ sends cokernels to kernels. However, Hom is *not* right exact in general.

Example 3.27. Consider the short exact sequence of abelian groups

$$0 \longrightarrow \mathbb{Z} \longrightarrow \mathbb{Q} \longrightarrow \mathbb{Q}/\mathbb{Z} \longrightarrow 0,$$

where the first map is the inclusion of $\mathbb{Z}$ into $\mathbb{Q}$, and the second map is the canonical projection. The elements in the abelian group $\mathbb{Q}/\mathbb{Z}$ are cosets of the form $\frac{p}{q} + \mathbb{Z}$, where $\frac{p}{q} \in \mathbb{Q}$, and whenever $\frac{p}{q} \in \mathbb{Z}$, $\frac{p}{q} + \mathbb{Z} = 0$. While Theorem 3.26 says that

$$0 \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2, \mathbb{Z}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2, \mathbb{Q}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2, \mathbb{Q}/\mathbb{Z})$$

is exact, we claim that this cannot be extended to a short exact sequence, since the map $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2, \mathbb{Q}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2, \mathbb{Q}/\mathbb{Z})$ is not surjective.

On the one hand, there are no nontrivial homomorphisms from $\mathbb{Z}/2$ to either $\mathbb{Z}$ nor $\mathbb{Q}$, since there are no elements in $\mathbb{Z}$ nor $\mathbb{Q}$ of order 2. This shows that

$$\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2, \mathbb{Q}) \cong 0.$$

On the other hand, $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2, \mathbb{Q}/\mathbb{Z})$ is nonzero: to give a homomorphism of abelian groups $\mathbb{Z}/2 \to \mathbb{Q}/ZZ$ is to choose an element in $\mathbb{Q}/\mathbb{Z}$ of order 2. Since $\frac{1}{2} + \mathbb{Z}$ is an element of order 2 in $\mathbb{Q}/\mathbb{Z}$, the map sending 1 in $\mathbb{Z}/2$ to $\frac{1}{2} + \mathbb{Z}$ in $\mathbb{Z}/\mathbb{Q}$ is nonzero. So after applying $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2, -)$, we get the exact sequence

$$0 \longrightarrow 0 \longrightarrow 0 \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2, \mathbb{Q}/\mathbb{Z}).$$

So this shows that $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2, -)$ is not an exact functor, only left exact.

Similarly, we can show that $\mathrm{Hom}_{\mathbb{Z}}(-, \mathbb{Z})$ is not exact:

Example 3.28. Let's apply $\mathrm{Hom}_{\mathbb{Z}}(-, \mathbb{Z})$ to the short exact sequence

$$0 \longrightarrow \mathbb{Z} \longrightarrow \mathbb{Q} \longrightarrow \mathbb{Q}/\mathbb{Z} \longrightarrow 0.$$

This time, Theorem 3.26 says that

$$0 \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q}/\mathbb{Z}, \mathbb{Z}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q}, \mathbb{Z}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}, \mathbb{Z}).$$

is exact. We claim that the last map is not surjective.

First, we claim that $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q}, \mathbb{Z}) = 0$. Indeed, if $f : \mathbb{Q} \longrightarrow \mathbb{Z}$ is a homomorphism of abelian groups, then for all $n \geqslant 1$ we have

$$f(1) = nf(\frac{1}{n}).$$

So $f(1)$ is an integer that is divisible by every integer, which is impossible unless $f(1) = 0$. We conclude that $f = 0$, and thus $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q}, \mathbb{Z}) \cong 0$. So our exact sequence above is actually

$$0 \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q}/\mathbb{Z}, \mathbb{Z}) \longrightarrow 0 \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}, \mathbb{Z}).$$

By Theorem 3.4, $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z},\mathbb{Z}) \cong \mathbb{Z} \neq 0$, so the last map in our sequence can't possibly be surjective, so our sequence is not a short exact sequence.

The other fun consequence is that since $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q},\mathbb{Z}) = 0$ and we have an exact sequence

$$0 \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q}/\mathbb{Z},\mathbb{Z}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q},\mathbb{Z}) = 0,$$

we can now conclude that

$$\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Q}/\mathbb{Z},\mathbb{Z}) = 0.$$

The last observation is a common trick: once we know we have an exact sequence involving certain modules we do not know, we can sometimes calculate them exactly by studying the other modules and maps in the exact sequence.

We can use the left exactness of Hom to compute some modules of interest:

Example 3.29. Let R be a commutative ring and M be a finitely presented R-module. This means that M has a presentation with finitely many generators and relations, which translates into an exact sequence of the form

$$R^m \xrightarrow{f} R^n \longrightarrow M \longrightarrow 0.$$

Since R^m and R^n are free modules, we can think of the map f as multiplication by a matrix A with n rows and m columns, after we fix a basis for R^n and R^m. Applying $\mathrm{Hom}_R(-,R)$ to the exact sequence above, we get an exact sequence

$$0 \longrightarrow \mathrm{Hom}_R(M,R) \longrightarrow \mathrm{Hom}_R(R^n,R) \xrightarrow{f^*} \mathrm{Hom}(R^m,R).$$

By Theorem 3.4, $\mathrm{Hom}_R(R^n,R) \cong R^n$ and $\mathrm{Hom}_R(R^m,R) \cong R^m$. Moreover, we claim that f^* is multiplication by the transpose of A.

First, note that given a basis $\{e_1,\dots,e_n\}$ for R^n, we get a dual basis $\{e_1^*,\dots,e_n^*\}$ for $\mathrm{Hom}_R(R^n,R)$, where

$$e_i^*(e_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, we have a dual basis $\{e_1^*,\dots,e_m^*\}$ for $\mathrm{Hom}_R(R^m,R) \cong R^m$; we might as well assume that we picked the canonical basis in both cases, so that we can use similar notation on both.

Now the map f^* is also given by multiplication by a matrix, now having m rows and n columns. To calculate its jth column, we need to calculate $f^*(e_j^*)$, which is given by precomposition with f, so $f^*(e_j^*) = e_j^* A$; this reads off the jth row of A. Thus f^* is indeed multiplication by A^T, and we have an exact sequence

$$0 \longrightarrow \mathrm{Hom}_R(M,R) \longrightarrow R^n \xrightarrow{A^T} R^m.$$

In particular, we have shown that $\mathrm{Hom}_R(M,R)$ is the kernel of multiplication by A^T.

3.2 Tensor products

Definition 3.30. Fix a ring R, and consider:

- a right R-module M,
- a left R-module N,
- an abelian group L.

A function $f\colon M \times N \longrightarrow L$ is **R-biadditive** if for all $m, m' \in M$, all $n, n' \in N$, and all $r \in R$ we have

- $f(m + m', n) = f(m, n) + f(m', n)$
- $f(m, n + n') = f(m, n) + f(m, n')$
- $f(mr, n) = f(m, rn)$.

When R is a commutative ring, suppose that L is also an R-module. We say that a function $f\colon M \times N \longrightarrow L$ is **R-bilinear** if for all $m, m' \in M$, all $n, n' \in N$, and all $r \in R$ we have

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

Note that an R-bilinear function is an R-biadditive function that satisfies

$$f(rm, n) = f(m, rn) = rf(m, n).$$

Example 3.31. The product on R is an R-biadditive function $R \times R \longrightarrow R$. The first two rules follow from distributivity of multiplication over the sum; the final rule is a consequence of the associativity of multiplication.

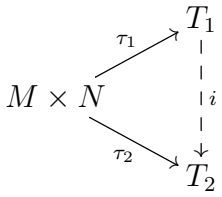
When R is commutative, this is an R-bilinear function.

Definition 3.32. Let M be a right R-module and let N be a left R-module. The **tensor product** of M and N is an abelian group $M \otimes_R N$ together with an R-biadditive function $\tau\colon M \times N \longrightarrow M \otimes_R N$ with the following universal property: for every abelian group A and every R-biadditive map $f\colon M \times N \longrightarrow A$, there exists a unique group homomorphism $\tilde{f}\colon M \otimes_R N \longrightarrow A$ such that the following diagram commutes:

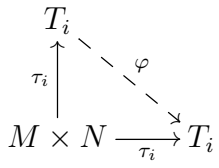
$$\begin{array}{ccc} M \otimes_R N & & \\ {\scriptstyle\tau}\uparrow & \searrow^{\tilde{f}} & \\ M \times N & \xrightarrow[f]{} & A \end{array}$$

We will now show that tensor products exist and are unique up to isomorphism; in particular, we can talk about *the* tensor product of M and N.

Lemma 3.33. *Let R be any ring, M be a right R-module, and N a left R-module. The tensor product of M and N is unique up to unique isomorphism. More precisely, if $M \times N \xrightarrow{\tau_1} T_1$ and $M \times N \xrightarrow{\tau_2} T_2$ are two tensor products, then there exists a unique isomorphism $T_1 \xrightarrow{i} T_2$ such that*

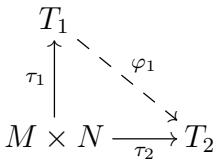


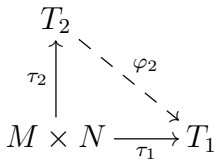
Proof. First, note that the universal property of the tensor product implies that there exists a unique φ such that



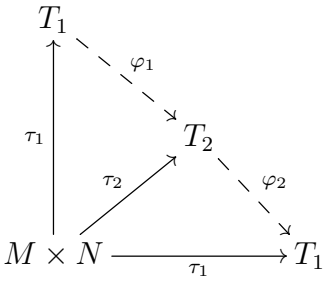
commutes. Since the identity map $T_i \longrightarrow T_i$ *is* such a map, it must be the *only* such map.

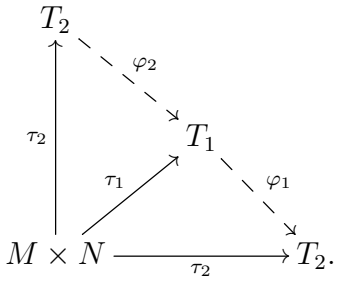
Similarly, there are unique maps $\varphi_1 \colon T_1 \longrightarrow T_2$ and $\varphi_2 \colon T_2 \longrightarrow T_1$ such that



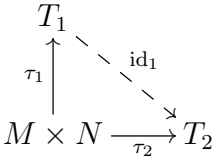


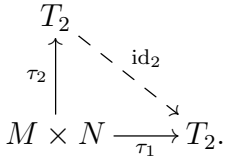
both commute. Stacking these up, we get commutative diagrams





Note that the identity maps on T_1 and T_2 *are* homomorphisms $T_1 \to T_1$ and $T_2 \to T_2$ that would make each of these triangles commute:





By uniqueness, $\varphi_2\varphi_1$ must be the identity on T_1 and $\varphi_1\varphi_2$ must be the identity on T_2. In particular, T_1 and T_2 are isomorphic, and the isomorphisms φ_1 and φ_2 are unique. $\square$

Theorem 3.34. *Given any right R-modules M and any left R-module N, their tensor product $M \otimes_R N$ exists, and it is given by the abelian group $M \otimes_R N$ defined as follows:*

- *Generators: For each pair of elements $m \in M$ and $n \in N$, we have a generator $m \otimes n$.*
- *Relations: the generators of $m \otimes n$ satisfy the following relations, where $m, m' \in M$, $n, n' \in N$, and $r \in R$:*

$$\begin{aligned} m \otimes (n + n') &= m \otimes n + m \otimes n' \\ (m + m') \otimes n &= m \otimes n + m \otimes n' \\ (mr) \otimes n &= m \otimes (rn). \end{aligned}$$

Proof. Let F be the free abelian group on the set $M \times N$. In what follows, we identify a pair $(m, n) \in M \times N$ with the corresponding basis element for F. Let S be the subgroup of F generated by

$$S = \left(\left\{ \begin{array}{c|c} (m, n + n') - (m, n) - (m, n') & m, m' \in M \\ (m + m', n) - (m, n) - (m', n) & n, n' \in N \\ (mr, n) - (m, rn) & r \in R \end{array} \right\} \right).$$

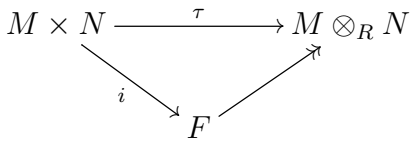
Let $M \otimes_R N := F/S$, and let $m \otimes n$ denote the class of (m, n) in the quotient. We claim that this abelian group $M \otimes_R N$ is a tensor product for M and N, together with the map

$$\begin{aligned} M \times N &\xrightarrow{\tau} M \otimes N \\ (m, n) &\longmapsto m \otimes n \end{aligned}$$

Notice τ is the restriction of the quotient map $F \longrightarrow F/S$ to the basis elements of F. Moreover, by construction of $M \otimes_R N$, the following identities hold:

$$\begin{aligned} m \otimes (n + n') &= m \otimes n + m \otimes n' \\ (m + m') \otimes n &= m \otimes n + m \otimes n' \\ (mr) \otimes n &= m \otimes (rn) \end{aligned}$$

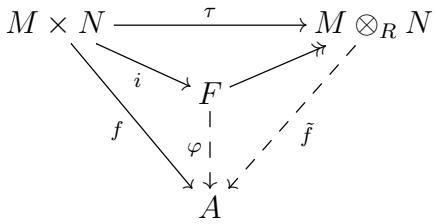
Together, these make τ an R-biadditive map. The map $M \times N \longrightarrow F$ that sends each pair (m, n) to the corresponding basis element is R-bilinear by construction. Moreover, there is a natural quotient map $F \longrightarrow M \otimes_R N$, and these maps make the diagram



commute.

Now suppose that A is any other abelian group, and let $M \times N \xrightarrow{f} A$ by any R-biadditive map. Since F is the free R-module on $M \times N$, f induces a homomorphism of abelian groups $\varphi \colon F \longrightarrow A$ such that $\varphi \circ i = f$, meaning $f(m, n) = \varphi(m, n)$ for all $m \in M$ and all $n \in N$.

Finally, the fact that f is bilinear implies that $S \subseteq \ker \varphi$. Therefore, φ induces a group homomorphism on $F/S = M \otimes_R N$. All this fits in the following commutative diagram:



Finally, this map $\tilde{f}$ we constructed satisfies $\tilde{f}(n \otimes n) = f(m, n)$, and since $M \otimes_R N$ is generated by such elements, $\tilde{f}$ is completely determined by the images of $m \otimes n$, and thus unique. □

The construction in Theorem 3.34 gives us generators $m \otimes n$ for $M \otimes_R N$. These are usually called **simple tensors**. So any element in $M \otimes_R N$ is of the form

$$\sum_{i=1}^{k} m_i \otimes n_i.$$

Such expressions are *not* unique. For a cheap example, consider the relations we used to construct $M \otimes_R N$ from the abelian group on $M \times N$, which gives us nontrivial ways to write the 0 element in $M \otimes_R N$:

$$\begin{aligned} 0 &= m \otimes (n + n') - m \otimes n - m \otimes n' \\ 0 &= (m + m') \otimes n - m \otimes n - m \otimes n' \\ 0 &= (mr) \otimes n - m \otimes (rn). \end{aligned}$$

This makes things unexpectedly tricky. For starters, the tensor product of two nonzero modules might be zero nevertheless. Also, whenever we try to define some R-module homomorphism from $M \otimes_R N$ into some other R-module, we must carefully check that our map is well-defined, which is in principle not an easy task. Therefore, the easiest way to define some R-module homomorphism from $M \otimes_R N$ is to give some R-bilinear map from $M \times N$ into our desired R-module.

In summary: the tensor product $M \otimes_R N$ of M and N is generated by the simple tensors $m \otimes n$, but it's important to remember (though we're all bound to forget once or twice) that *not* all elements in $M \otimes_R N$ are simple tensors. Moreover, even if M and N are nonzero, $M \otimes_R N$ could very well be zero.

Remark 3.35. Two group homomorphisms $M \otimes_R N \longrightarrow L$ coincide if and only if they agree on simple tensors, since these are generators for $M \otimes_R N$.

Remark 3.36. In any tensor product $M \otimes_R N$, the simple tensor $0 \otimes 0$ is the zero element, and

$$m \otimes 0 = 0 = 0 \otimes n$$

for all $m \in M$ and $n \in N$.

Let's see some examples of how tensor products can be zero.

Example 3.37. We claim that $\mathbb{Z}/2 \otimes_{\mathbb{Z}} \mathbb{Q} = 0$, despite the fact that both of these $\mathbb{Z}$-modules are nonzero. To see that, simply note that given any $a \in \mathbb{Z}/2$ and any $p \in \mathbb{Q}$,

$$a \otimes p = a \otimes \frac{2p}{2} = (2a) \otimes \frac{p}{2} = 0 \otimes \frac{p}{2} = 0.$$

Since $\mathbb{Z}/2 \otimes_{\mathbb{Z}} \mathbb{Q}$ is generated by simple tensors, which are all 0, we conclude that $\mathbb{Z}/2 \otimes_{\mathbb{Z}} \mathbb{Q} = 0$.

Example 3.38. Consider the abelian group $\mathbb{Q}/\mathbb{Z}$. Again, this is very much nonzero, and yet we claim that $\mathbb{Q}/\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Q}/\mathbb{Z} = 0$. For any simple tensor,

$$\begin{aligned}\left(\frac{p}{q} + \mathbb{Z}\right) \otimes \left(\frac{a}{b} + \mathbb{Z}\right) = \left(\frac{bp}{bq} + \mathbb{Z}\right) \otimes \left(\frac{a}{b} + \mathbb{Z}\right) &= \left(\frac{p}{bq} + \mathbb{Z}\right) \otimes b\left(\frac{a}{b} + \mathbb{Z}\right) \\ &= \left(\frac{p}{bq} + \mathbb{Z}\right) \otimes 0 = 0 \otimes 0 = 0.\end{aligned}$$

Example 3.39. Let p and q be distinct prime integers. Then p has inverse modulo q, say $ap \equiv 1 \bmod q$, and q has an inverse modulo p, say $bq \equiv 1 \bmod p$. Given any simple tensor $n \otimes m$ in $\mathbb{Z}/p \otimes_{\mathbb{Z}} \mathbb{Z}/q$,

$$n \otimes m = ((bq)n) \otimes ((ap)m) = (pbn) \otimes (qam) = 0 \otimes 0.$$

Since all simple tensors are 0 and $\mathbb{Z}/p \otimes_{\mathbb{Z}} \mathbb{Z}/q$ is generated by simple tensors, we conclude that $\mathbb{Z}/p \otimes_{\mathbb{Z}} \mathbb{Z}/q = 0$.

More generally, the following holds:

Exercise 3.40. Show that if $d = \gcd(m, n)$, then $\mathbb{Z}/n \otimes_{\mathbb{Z}} \mathbb{Z}/m \cong \mathbb{Z}/d$.

Of course not all tensor products are zero. A good method for showing that a particular element m in a module M is nonzero is to give a homomorphism from M sending m to some nonzero element. We apply this technique to tensor products: to show that a particular element x in $M \otimes_R N$ is nonzero, we construct a homomorphism from $M \otimes_R N$ that takes x no some nonzero element. This is typically easier for simple tensors: we need an R-biadditive map out of $M \times N$ that sends the corresponding pair to a nonzero element.

Example 3.41. Consider the abelian group $2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/2$. The map

$$\begin{aligned}2\mathbb{Z} \times \mathbb{Z}/2 &\longrightarrow \mathbb{Z}/2 \\ (a, b) &\longmapsto \tfrac{ab}{2}\end{aligned}$$

is $\mathbb{Z}$-bilinear, and thus it induces a homomorphism $2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/2 \longrightarrow \mathbb{Z}/2$. Via this map, $2 \otimes 1 \mapsto 1 \neq 0$, so $2 \otimes 1$ is nonzero in $2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/2$, and $2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/2 \neq 0$.

Moreover, not all elements in a tensor product are simple tensors.

Exercise 3.42. Let $R = \mathbb{Z}[x]$ and consider the ideal $I = (2, x)$. Show that in $I \otimes_R I$, the element $2 \otimes 2 + x \otimes x$ is not a simple tensor.

We can sometimes give $M \otimes_R N$ the structure of an R-module.

Remark 3.43. Let R be a commutative ring, and let M and N be R-modules. We can give $M \otimes_R N$ the structure of an R-module, as follows: given $r \in R$ and a simple tensor $m \otimes n$,

$$r(m \otimes n) = (rm) \otimes n = m \otimes (rn).$$

We can then extend this linearly to all other elements of $M \otimes_R N$. We leave it as an exercise to check that this does indeed make the abelian group $M \otimes_R N$ into an R-module.

Alternatively, over a commutative ring we can define the tensor product as follows:

Definition 3.44. Let R be a commutative ring and M and N be R-modules. The **tensor product** of M and N is an R-module $M \otimes_R N$ together with an R-bilinear map $\tau \colon M \times N \longrightarrow M \otimes_R N$ with the following universal property: for every R-module A and every R-bilinear map $f \colon M \times N \longrightarrow A$ there exists a unique R-module homomorphism $\tilde{f} \colon M \otimes_R N \longrightarrow A$ such that the following diagram commutes:

$$\begin{array}{ccc} M \otimes_R N & & \\ {\scriptstyle\tau}\uparrow & \searrow^{\tilde{f}} & \\ M \times N & \xrightarrow[f]{} & A \end{array}$$

One can now check that if we take the abelian group $M \otimes_R N$, which is the unique abelian group which satisfies the universal property of the tensor product (as defined for a general ring R), and endow it with the R-module structure defined in Theorem 3.43, the resulting R-module satisfies the universal property in Theorem 3.44, and the argument we gave in Theorem 3.33 can be repurposed to show that this is the unique R-module satisfying this universal property.

Remark 3.45. We can express the universal property of the tensor product in the framework of Theorem 1.109. For simplicity, assume that R is a commutative ring. Consider the functor $\mathrm{Bilin}(M \times N, -) \colon R\textbf{-Mod} \longrightarrow \textbf{Set}$ that sends an R-module A to the set of R-bilinear maps $M \times N \longrightarrow A$, and a map of R-modules $f\ A \longrightarrow B$ to the function of sets induced by post-composition of functions. The universal property of the tensor product is encoded in the representable functor $\mathrm{Bilin}(M \times N, -) : R\textbf{-Mod} \longrightarrow \textbf{Set}$ together with the bilinear map $\tau \in \mathrm{Bilin}(M \times N, M \otimes_R N)$. Indeed, this says that τ induces a natural isomorphism between $\mathrm{Hom}_R(M \otimes_R N, -)$ and $\mathrm{Bilin}(M \times N, -)$ by sending each R-module A to the bijection

$$\begin{aligned} \mathrm{Hom}_R(M \otimes_R N, A) &\longrightarrow \mathrm{Bilin}(M \times N, A) \\ f &\longmapsto \mathrm{Bilin}(M \times N, f)\tau = f_*(\tau) = f\tau. \end{aligned}$$

The fact that this is a bijection says that for every R-bilinear map g there exists a unique R-module homomorphism f such that

$$\begin{array}{ccc} M \otimes_R N & & \\ {\scriptstyle\tau}\uparrow & \searrow^{f} & \\ M \times N & \xrightarrow[g]{} & A \end{array}$$

commutes. So this is indeed the universal property we described before.

More generally, $M \otimes_R N$ has a module structure when one of M or N is a bimodule.

Definition 3.46. Fix rings R and S. An (R, S)**-bimodule** is an abelian group M together with a left R-module structure and a right S-module structure such that for all $r \in R$, $s \in S$, and $m \in M$,

$$(rm)s = r(ms).$$

One sometimes writes ${}_RM_S$ to indicate M is an (R, S)-bimodule. An R**-bimodule** is an (R, R)-bimodule.

Example 3.47.

a) Let $\mathrm{M}_{m,n}(R)$ denote the ring of $m \times n$ matrices with entries in a ring R. We can also view $\mathrm{M}_{m,n}(R)$ as an $(\mathrm{M}_{m,m}, \mathrm{M}_{n,n})$-bimodule via left and right multiplication of matrices.

b) Any two-sided ideal I of a ring R is an R-bimodule.

c) Let R be a commutative ring and let M be any left R-module. Then M is also a right R-module under the same module structure, by setting

$$m \cdot r := rm.$$

Moreover, M is also an R-bimodule using both of these structures at once.

d) Let $f\colon R \to S$ be a ring homomorphism. We can view S as an (R, S)-bimodule via

$$t \cdot s \cdot r := tsf(r)$$

for $t, s \in S$ and $r \in R$, where the right hand side is just multiplication in s. Similarly, S can be viewed as an (S, R)-bimodule and as an (R, R)-bimodule.

e) Let R be a commutative ring of prime characteristic $p > 0$, meaning that R contains a copy of $\mathbb{F}_p$, or equivalently, that

$$\underbrace{1 + \cdots + 1}_{p \text{ times}} = 0.$$

Then R is an R-bimodule with the left module structure given by the Frobenius map

$$\begin{array}{ccc} R & \xrightarrow{F} & R \\ r & \longmapsto & r^p \end{array}$$

and right module structure given by the usual multiplication on R. More precisely, given $r, s, t \in R$,

$$r \cdot s \cdot t := r^p st$$

where the right hand side is just multiplication in R.

Exercise 3.48. Let M be an (S, R)-bimodule and N a left R-module. Consider $M \times N$ as a left S-module via

$$s(m, n) = (sm, n).$$

Then $M \otimes_R N$ is a left S-module via

$$s\left(\sum_i m_i \otimes n_i\right) = (sm_i) \otimes n_i.$$

The map

$$\begin{aligned} M \times N &\longrightarrow M \otimes_R N \\ (m, n) &\longrightarrow m \otimes n \end{aligned}$$

is left S-linear, and for any left S-module A and left S-linear R-biadditive map $b\colon M \times N \to A$, there is a unique left S-linear map $\alpha\colon M \otimes_R N \to A$ such that $\alpha(m \otimes n) = b(m, n)$.

Similarly, for a left R-module M and an (R, S)-bimodule N, $M \times N$ is a right S-module via

$$(m, n)s = (m, ns).$$

Then $M \otimes_R N$ is a right S-module via

$$\left(\sum_i m_i \otimes n_i\right) s = m_i \otimes (n_i s),$$

and the map

$$\begin{aligned} M \times N &\longrightarrow M \otimes_R N \\ (m, n) &\longrightarrow m \otimes n \end{aligned}$$

is right S-linear, and for any S-module A and right S-linear R-biadditive map $b\colon M \times N \to A$, there is a unique right S-linear map $\alpha\colon M \otimes_R N \to A$ such that $\alpha(m \otimes n) = b(m, n)$.

We can also take tensor products of maps.

Lemma 3.49. *Let R be a ring, $f : A \to C$ be a homomorphism of right R-modules, and $g\colon B \to D$ be a homomorphism of left R-modules. There exists a unique homomorphism of abelian groups $f \otimes g\colon A \otimes_R B \longrightarrow C \otimes_R D$ such that*

$$(f \otimes g)(a \otimes b) = f(a) \otimes g(b)$$

for all $a \in A$ and $b \in B$. When R is commutative, this map $f \otimes g$ is a homomorphism of R-modules. Moreover, if A and B are (S, R)-bimodules and f is left S-linear, then $f \otimes g$ is also a homomorphism of left S-modules, and if C and D are (R, S)-bimodules and g is right S-linear, then $f \otimes g$ is also a homomorphism of right S-modules.

Proof sketch. The function

$$\begin{aligned} A \times B &\longrightarrow C \otimes_R D \\ (a, b) &\longmapsto f(a) \otimes g(b) \end{aligned}$$

is R-biadditive, and R-bilinear when R is commutative, and right or left S-linear in the bimodule case, so the universal property of tensor products in each case gives the desired homomorphism and its uniqueness. □

Lemma 3.50. *Given R-module maps $A_1 \xrightarrow{f_1} A_2 \xrightarrow{f_2} A_3$ and $B_1 \xrightarrow{b_1} B_2 \xrightarrow{g_2} B_3$, the composition of $f_1 \otimes g_1$ satisfies $f_2 \otimes g_2$*

$$(f_2 \otimes g_2) \circ (f_1 \otimes g_1) = (f_2 f_1) \otimes (g_2 g_1).$$

Proof. It's sufficient to check that these maps agree on simple tensors, and indeed they both take $a \otimes b$ to $(f_2 f_1(a)) \otimes (g_2 g_1(b))$. □

We are particularly interested in tensor products because of the tensor functor.

Theorem 3.51. *Let M be a right R-module. There is an additive covariant functor*

$$M \otimes_R -\colon R\textbf{-}\boldsymbol{Mod} \longrightarrow \mathbf{Ab}$$

that takes each R-module N to $M \otimes_R N$, and each R-module homomorphism $f\colon A \longrightarrow B$ to the homomorphism of abelian groups $1_M \otimes f\colon M \otimes_R A \longrightarrow M \otimes_R B$.

When R is commutative, we can view $M \otimes_R -$ as an additive functor $R\textbf{-}\boldsymbol{Mod} \to R\textbf{-}\boldsymbol{Mod}$.

Proof. Let $T := M \otimes_R -$. First, note that T preserves identities, meaning $T(1_N) = 1_{T(N)}$, since the identity map on $M \otimes_R N$ agrees with $T(1_N) = 1_M \otimes 1_N$ on simple tensors. Moreover, T preserves compositions, since by Theorem 3.50 we have

$$T(f)T(g) = (1 \otimes f)(1 \otimes g) = 1 \otimes (fg) = T(fg).$$

Therefore, T is a functor. To check that it is an additive functor, we need to prove that $T(f+g) = T(f) + T(g)$ for all $f, g \in \mathrm{Hom}_R(A, B)$. It is sufficient to check that the maps $T(f+g) = 1 \otimes (f+g)$ and $T(f) + T(g) = 1 \otimes f + 1 \otimes g$ agree on simple tensors. Indeed,

$$\begin{aligned} T(f+g)(a \otimes b) &= (1 \otimes (f+g))(a \otimes b) \\ &= a \otimes (f+g)(b) \\ &= a \otimes f(b) + g(b) \\ &= a \otimes f(b) + a \otimes g(b) \\ &= (1 \otimes f)(a \otimes b) + (1 \otimes g)(a \otimes b) \\ &= T(f)(a \otimes b) + T(g)(a \otimes b). \end{aligned}$$

We conclude that $T(f+g) = T(f) + T(g)$. □

Definition 3.52. Given a ring R and a right R-module M, the functor $M \otimes_R -$ is the **tensor product functor**.

Note that we were purposely vague on the target of the tensor product functor: when R is commutative, we get both a functor R-**Mod** $\to$ **Ab** and a functor R-**Mod** $\to R$-**Mod**. The two functors are essentially the same: the tensor product functor R-**Mod** $\to$ **Ab** is the composition of functor R-**Mod** $\to R$-**Mod** followed by the forgetful functor R-**Mod** $\to$ **Ab**.

We can similarly define the tensor product functor $- \otimes_R N$; when R is commutative, it turns out that the two constructions are essentially the same.

Lemma 3.53 (Commutativity of tensor products)**.** *Let R be a commutative ring. There is a natural isomorphism $M \otimes_R - \cong - \otimes_R N$. In particular, for all R-modules M and N we have*

$$M \otimes_R N \cong N \otimes_R M.$$

Proof. One can check (exercise!) that the map $M \times N \longrightarrow N \otimes_R M$ given by $(m, n) \mapsto n \otimes m$ is R-biadditive, and R-bilinear if R is commutative. The universal property of the tensor product $M \otimes_R N$ gives us a homomorphism φ of abelian groups or R-modules, depending on the case, such that the diagram

$$\begin{array}{ccc} & & M \otimes_R N \\ & \nearrow & \vdots \\ M \times N & \longrightarrow & N \otimes_R M \\ (m,n) & \longmapsto & n \otimes m \end{array} \qquad \begin{array}{c} \\ \\ \varphi \\ \\ \end{array}$$

commutes. Similarly, we get a map ψ and a commutative diagram

$$\begin{array}{ccc} & & N \otimes_R M \\ & \nearrow & \vdots \\ N \times M & \longrightarrow & M \otimes_R N \\ (m,n) & \longmapsto & n \otimes m \end{array} \qquad \begin{array}{c} \\ \\ \psi \\ \\ \end{array}$$

Then $\varphi\psi$ agrees with the identity on $N \otimes_R M$ on simple tensors, so it is the identity. Similarly, $\psi\varphi$ is the identity on $M \otimes_R N$, and these are the desired isomorphisms.

The statement about naturality is more precisely the following: for every R-module maps $f : M_1 \longrightarrow M_2$ and $g : N_1 \longrightarrow N_2$, our isomorphisms $M_1 \otimes_R N_1 \cong N_1 \otimes_R M_1$ and $M_2 \otimes_R N_2 \cong N_2 \otimes_R M_2$ make the diagram

$$\begin{array}{ccc} M_1 \otimes_R N_1 & \xrightarrow{\cong} & N_1 \otimes_R M_1 \\ \downarrow{\scriptstyle f \otimes g} & & \downarrow{\scriptstyle g \otimes f} \\ M_2 \otimes_R N_2 & \xrightarrow{\cong} & N_2 \otimes_R M_2 \end{array}$$

commute. To check this, it's sufficient to check commutativity on simple tensors, and indeed

$$\begin{array}{ccc} m \otimes n & \longmapsto & n \otimes m \\ \downarrow & & \downarrow \\ f(m) \otimes g(n) & \longmapsto & g(n) \otimes f(m). \end{array} \qquad \begin{array}{ccc} M_1 \otimes_R N_1 & \xrightarrow{\cong} & N_1 \otimes_R M_1 \\ \downarrow{\scriptstyle f \otimes g} & & \downarrow{\scriptstyle g \otimes f} \\ M_2 \otimes_R N_2 & \xrightarrow{\cong} & N_2 \otimes_R M_2 \end{array}$$

Lemma 3.54 (Associativity of tensors)**.** *Given a right R-module A, an (R, S)-bimodule B, and a left S-module C,*

$$(A \otimes_R B) \otimes_S C \cong A \otimes_R (B \otimes_S C).$$

Proof. Fix $c \in C$. The map

$$\begin{aligned} A \times B &\longrightarrow A \otimes_R (B \otimes_R C) \\ (a, b) &\longmapsto a \otimes (b \otimes c) \end{aligned}$$

is R-biadditive, so it induces a homomorphism of abelian groups

$$\varphi_c \colon A \otimes_R B \longrightarrow A \otimes_R (B \otimes_R C).$$

This map is in fact a homomorphism of R-modules when R is commutative. Moreover,

$$\begin{aligned} (A \otimes_R B) \times C &\longrightarrow A \otimes_R (B \otimes_R C) \\ (a \otimes b, c) &\longmapsto a \otimes (b \otimes c) \end{aligned}$$

is also R-biadditive, and it induces a homomorphism that sends $(a \otimes b) \otimes c$ to $a \otimes (b \otimes c)$. Similarly, we can define a homomorphism

$$\begin{aligned} A \otimes_R (B \otimes_R C) &\longrightarrow (A \otimes_R B) \otimes_R C \\ a \otimes (b \otimes c) &\longmapsto (a \otimes b) \otimes c. \end{aligned}$$

The composition of these two homomorphisms in either order is the identity on simple tensors, and thus they are both isomorphisms. □

Lemma 3.55. *Let R be any ring. There is a natural isomorphism between $R \otimes_R -$ and the identity functor on R-**Mod**. In particular, for every left R-module M there is an isomorphism of R-modules*

$$R \otimes_R M \cong M.$$

Proof. First, note that R is an R-bimodule, so $R \otimes_R M$ is a left R-module. The map

$$\begin{aligned} R \times M &\longrightarrow M \\ (r, m) &\longmapsto rm \end{aligned}$$

is R-biadditive (by the distributive laws), R-bilinear (by associativity of the action on a module), and R-linear, so it induces a homomorphism of R-modules $R \otimes_R M \xrightarrow{\varphi_M} M$. By definition, φ_M is surjective. Moreover, the map

$$\begin{aligned} M &\xrightarrow{f_M} R \otimes_R M \\ m &\longmapsto 1 \otimes m \end{aligned}$$

is a homomorphism of R-modules, since

$$f_M(a + b) = 1 \otimes (a + b) = 1 \otimes a + 1 \otimes b \text{ and } f_M(ra) = 1 \otimes (ra) = r(1 \otimes a) = r f_M(a).$$

For every $m \in M$, $\varphi_M f_M(m) = \varphi_M(1 \otimes m) = 1m = m$, and for every simple tensor, $f_M \varphi_M(r \otimes m) = f_M(rm) = 1 \otimes (rm) = r \otimes m$. This shows that φ_M is an isomorphism.

Finally, given any $f \in \mathrm{Hom}_R(M, N)$, since f is R-linear we conclude that the diagram

$$\begin{array}{ccc}
r \otimes m & \longmapsto & rm \\
\Big\downarrow & & \Big\downarrow \\
R \otimes_R M & \xrightarrow{\varphi_M} & M \\
{\scriptstyle 1\otimes f}\Big\downarrow & & \Big\downarrow{\scriptstyle f} \\
R \otimes N & \xrightarrow[\varphi_N]{} & N \\
\Big\downarrow & & \Big\downarrow \\
r \otimes f(m) & \longmapsto & rf(m) = f(rm)
\end{array}$$

commutes, so our isomorphism is natural. □

Similarly to the Hom functor, tensor behaves well with respect to arbitrary direct sums.

Theorem 3.56. *Let M be a right R-module, and let $\{N_i\}_{i \in I}$ be an arbitrary family of left R-modules. Then the map*

$$\begin{aligned}
M \otimes_R \left(\bigoplus_{i \in I} N_i \right) &\overset{\cong}{\to} \bigoplus_{i \in I} M \otimes_R N_i \\
m \otimes (a_i)_i &\longmapsto (m \otimes a_i)
\end{aligned}$$

is an isomorphism of abelian groups in general, of R-modules in the commutative case, of S-modules if each N_i is an (S, R)-bimodule, and of right S-modules if N is an (R, S)-bimodule. Moreover, this isomorphism is natural: given two families of left R-modules $\{A_i\}_{i \in I}$ and $\{B_j\}_{j \in J}$, and left R-module homomorphisms $\sigma_{ij} \colon A_i \longrightarrow B_j$, the R-module homomorphisms

$$\begin{aligned}
&\bigoplus_{i \in I} A_i \xrightarrow{\sigma} \bigoplus_{j \in J} B_j \qquad \text{and} \quad \tilde{\sigma} = \bigoplus_{i \in I} \sigma_{ij} \colon \bigoplus_{i \in I} M \otimes_R A_i \longrightarrow \bigoplus_{j \in J} M \otimes_R B_j \\
&(a_i)_{i \in I} \longmapsto (\sigma_{ij}(a_i))_{j \in J}
\end{aligned}$$

give a commutative diagram

$$\begin{array}{ccc}
M \otimes_R \left(\bigoplus_{i \in I} A_i \right) & \xrightarrow{\cong} & \bigoplus_{i \in I} M \otimes_R A_i \\
{\scriptstyle 1\otimes\sigma}\Big\downarrow & & \Big\downarrow{\scriptstyle \tilde{\sigma}} \\
M \otimes_R \left(\bigoplus_{j \in J} B_j \right) & \xrightarrow[\cong]{} & \bigoplus_{j \in J} M \otimes_R B_j.
\end{array}$$

Proof. First, note that the function

$$\begin{aligned} M \times \left(\bigoplus_{i \in I} A_i \right) &\longrightarrow \bigoplus_{i \in I} (M \otimes_R A_i) \\ (m, (a_i)_i) &\longmapsto (m \otimes a_i) \end{aligned}$$

is R-bilinear, so it induces a homomorphism

$$M \otimes_R \left(\bigoplus_{i \in I} A_i \right) \xrightarrow{\tau} \bigoplus_{i \in I} (M \otimes_R A_i) \ .$$

For each $k \in I$, let ι_k denote the inclusion map $A_k \subseteq \bigoplus_i A_i$. The universal property of the coproduct (which in the case of R-modules, means the direct sum) gives an R-module homomorphism

$$\begin{aligned} \bigoplus_{i \in I} (M \otimes_R A_i) &\xrightarrow{\lambda} M \otimes_R \bigoplus_{i \in I} (A_i) \\ (m \otimes a_i)_i &\longmapsto m \otimes \sum_i \iota_i(a_i) \end{aligned}$$

which we obtain by assembling the R-module homomorphisms $1 \otimes \iota_i$. It is routine to check that λ is the inverse of τ, which must then be an isomorphism. Finally, we can check naturality by checking commutativity of the square above, element by element:

$$\begin{array}{ccc} m \otimes (a_i)_i & \longmapsto & (m \otimes a_i)_i \\ \downarrow & & \downarrow \\ m \otimes (\sigma_{ij}(a_i))_i & \longmapsto & (m \otimes \sigma_{ij}(a_i)). \end{array}$$

Remark 3.57. By commutativity of the tensor product, we also get natural isomorphisms

$$\left(\bigoplus_{i \in I} N_i \right) \otimes_R M \xrightarrow{\cong} \bigoplus_{i \in I} N_i \otimes_R M.$$

The following follows as a corollary of Theorem 3.55 and Theorem 3.56:

Exercise 3.58. Show that if F and G are free R-modules on bases $\{e_\lambda\}_{\lambda \in \Lambda}$ and $\{e_\gamma\}_{\gamma \in \Gamma}$, respectively, then $F \otimes_R G$ is the free R-module on basis

$$\{e_\lambda \otimes e_\gamma \mid \lambda \in \Lambda, \gamma \in \Gamma\}.$$

In particular,

$$R^n \otimes R^m \cong R^{nm}.$$

Example 3.59. Let R be any ring and consider $R^2 \otimes_R R^2$. Let $e_1 = (1,0) \in R^2$ and $e_2 = (0,1) \in R^2$. We claim that the element $e_1 \otimes e_2 + e_2 \otimes e_1$ is not a simple tensor. Suppose, by contradiction, that there exist $v, y \in R^2$ such that

$$e_1 \otimes e_2 + e_2 \otimes e_1 = v \otimes w.$$

Since $\{e_1, e_2\}$ is a basis for the free module R^2, we can write

$$v = v_1 e_1 + v_2 e_2 \qquad \text{and} \qquad w = w_1 e_1 + w_2 e_2.$$

Substituting above, we see that

$$\begin{aligned} v \otimes w &= (v_1 e_1 + v_2 e_2) \otimes (w_1 e_1 + w_2 e_2) \\ &= v_1 w_1 e_1 \otimes e_1 + v_1 w_2 e_1 \otimes e_2 + v_2 w_1 e_2 \otimes e_1 + v_2 w_2 e_2 \otimes e_2. \end{aligned}$$

But by Theorem 3.58, $\{e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_1, e_2 \otimes e_2\}$ is a basis for the free R-module $R^2 \otimes R^2 \cong R^4$, so we can now compare coefficients: since

$$e_1 \otimes e_2 + e_2 \otimes e_1 = v_1 w_1 e_1 \otimes e_1 + v_1 w_2 e_1 \otimes e_2 + v_2 w_1 e_2 \otimes e_1 + v_2 w_2 e_2 \otimes e_2,$$

we must have

$$\begin{cases} v_1 w_1 = 1 \\ v_1 w_2 = 0 \\ v_2 w_1 = 0 \\ v_2 w_2 = 1 \end{cases} \implies \begin{cases} v_1 \text{ and } w_1 \text{ are units} \\ v_1 w_2 = 0 \\ v_2 w_1 = 0 \\ v_2 \text{ and } w_2 \text{ are units} \end{cases}$$

But since v_1 is a unit and $v_1 w_2 = 0$, we must have $w_2 = 0$; similarly, since v_2 is a unit and $v_2 w_1 = 0$, we must have $w_1 = 0$. But we have both $w_1 = w_2 = 0$ and that w_1, w_2 are units, which is a contradiction. We conclude that $e_1 \otimes e_2 + e_2 \otimes e_1$ is not a simple tensor.

One of the reasons tensor products are useful is that we can use tensor products to extend module structures to ring extensions.

Remark 3.60. Let $f : R \to S$ be a ring homomorphism. Since S is an (S, R)-bimodule, the abelian group $S \otimes_R M$ has a left S-module structure for every left R-module M. Thus $S \otimes_R -$ determines a functor from R-modules to S-modules.

Definition 3.61. Let $f : R \to S$ be a ring homomorphism. The **extension of scalars** from R to S is the functor $S \otimes_R - : R\textbf{-Mod} \longrightarrow S\textbf{-mod}$: for each R-module M, we get an S-module $S \otimes_R M$ with

$$s \cdot \left(\sum_i s_i \otimes m_i \right) := \sum_i (s s_i) \otimes m_i,$$

and for each R-module homomorphism $f : M \to N$ we get the S-module homomorphism $1 \otimes_R f : S \otimes_R M \longrightarrow S \otimes_R N$.

This functor is closely related to restriction of scalars: we will soon show that restriction and extension of scalars are adjoint functors.

Definition 3.62. Let $f : R \to S$ be a ring homomorphism. The **restriction of scalars functor** from S to R is the functor $f^* : S\textbf{-mod} \longrightarrow R\textbf{-Mod}$ that takes each S-module M to the R-module f^*M with underlying abelian group M and R-module structure

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

induced by f. Moreover, for each S-module homomorphism $g : M \longrightarrow N$ we get the R-module homomorphism $f^*(g) : f^*(M) \longrightarrow f^*(N)$ defined by $f^*(g)(m) := g(n)$.

Exercise 3.63. Check that restriction of scalars as defined above is indeed a functor.

Tensor is right exact.

Theorem 3.64. *Let M be a right R-module. The functor $M \otimes_R -$ is right exact, meaning that for every exact sequence*

$$A \xrightarrow{\ i\ } B \xrightarrow{\ p\ } C \longrightarrow 0$$

the sequence

$$M \otimes_R A \xrightarrow{1\otimes i} M \otimes_R B \xrightarrow{1\otimes p} M \otimes_R C \longrightarrow 0$$

is exact.

Proof. Since additive functors send complexes to complexes, $(1 \otimes p)(1 \otimes i) = 0$. We have two more things to show:

$\underline{1 \otimes p \text{ is surjective}}$: Consider any $m_1 \otimes c_1 + \cdots + m_n \otimes c_n \in M \otimes_R C$. Since p is surjective, we can find $b_1, \ldots, b_n \in B$ such that $p(b_i) = c_i$. Therefore,

$$(1 \otimes p)(m_1 \otimes b_1 + \cdots + m_n \otimes b_n) = m_1 \otimes p(b_1) + \cdots + m_n \otimes p(b_n) = m_1 \otimes c_1 + \cdots + m_n \otimes c_n.$$

$\underline{\ker(1 \otimes p) = \operatorname{im}(1 \otimes i)}$: Let $I = \operatorname{im}(1 \otimes i)$. We have already shown that $I \subseteq \ker(1 \otimes p)$, so $1 \otimes p$ induces a map $q : (M \otimes_R B)/I \longrightarrow M \otimes_R C$. Let $\pi : M \otimes_R B \longrightarrow (M \otimes_R B)/I$ be the canonical projection. By definition, $q\pi = 1 \otimes p$.

Consider the map

$$\begin{aligned} M \times C &\xrightarrow{\ f\ } (M \otimes_R B)/I\ , \\ (m, c) &\longmapsto m \otimes b \end{aligned}$$

where b is such that $p(b) = c$. First, we should check this map f is well-defined. To see that, suppose that $b' \in B$ is another element with $p(b') = c$, so that $p(b - b') = 0$. Then $b - b' \in \ker p = \operatorname{im} i$, so $m \otimes (b - b') \in \operatorname{im}(1 \otimes i) \subseteq I$. Therefore, $m \otimes b = m \otimes b'$ modulo I, and f is well-defined.

Moreover, one can check (exercise!) that f is R-biadditive, so it induces a homomorphism of R-modules $M \otimes_R C \longrightarrow (M \otimes_R B)/I$, which we will denote by $\hat{f}$. We will show that $\hat{f}$ is a left inverse of q, so q is injective. And indeed, given $m_i \in M$ and $b_i \in B$, we have

$$\hat{f}q\left(\sum_{i=1}^{n} m_i \otimes b_i\right) = f\left(\sum_{i=1}^{n} m_i \otimes p(b_i)\right) = \sum_{i=1}^{n} f(m_i \otimes p(b_i)) = \sum_{i=1}^{n} m_i \otimes b_i.$$

We conclude that q is injective, and thus

$$\ker(1 \otimes p) = \ker(q\pi) = \ker \pi = I = \operatorname{im}(1 \otimes i). \qquad \square$$

However, tensor is not exact.

Example 3.65. Consider the short exact sequence

$$0 \longrightarrow \mathbb{Z} \xrightarrow{i} \mathbb{Q} \xrightarrow{p} \mathbb{Q}/\mathbb{Z} \longrightarrow 0.$$

Applying the functor $\mathbb{Z}/2 \otimes_{\mathbb{Z}} -$, we get an exact sequence

$$\mathbb{Z}/2 \otimes_{\mathbb{Z}} \mathbb{Z} \xrightarrow{1\otimes i} \mathbb{Z}/2 \otimes_{\mathbb{Z}} \mathbb{Q} \xrightarrow{1\otimes p} \mathbb{Z}/2 \otimes_{\mathbb{Z}} \mathbb{Q}/\mathbb{Z} \longrightarrow 0.$$

However, we claim that $1 \otimes i$ is not injective. On the one hand, by Theorem 3.55 we have an isomorphism $\mathbb{Z}/2 \otimes_{\mathbb{Z}} \mathbb{Z} \cong \mathbb{Z}/2 \neq 0$. On the other hand, we have seen in Theorem 3.37 that $\mathbb{Z}/2 \otimes_{\mathbb{Z}} \mathbb{Q} = 0$, so the map $1 \otimes i \colon \mathbb{Z}/2 \to 0$ cannot possibly be injective.

We can now show that extension of scalars turns an R-module into the S-module with the same presentation.

Remark 3.66. Let R be a ring, M be a right R-module, and N be a left R-module. We can compute $M \otimes_R N$ by taking a presentation of M

$$R^{\oplus\Gamma} \xrightarrow{\phi} R^{\oplus\Lambda} \longrightarrow M \longrightarrow 0$$

and tensoring with N to get

$$N^{\oplus\Gamma} \longrightarrow N^{\oplus\Lambda} \longrightarrow M \otimes_R N \longrightarrow 0,$$

so $M \otimes_R N$ is the cokernel of the map $N^{\oplus\Gamma} \to N^{\oplus\Lambda}$ induced by ϕ. We can also compute $M \otimes_R N$ by taking a presentation of N

$$R^{\oplus\Xi} \xrightarrow{\psi} R^{\oplus\Omega} \longrightarrow N \longrightarrow 0$$

and tensoring with M to get

$$M^{\oplus\Xi} \longrightarrow M^{\oplus\Omega} \longrightarrow M \otimes_R N \longrightarrow 0,$$

so $M \otimes_R N$ is isomorphic to the cokernel of the map $M^{\oplus\Gamma} \to M^{\oplus\Lambda}$ induced by ψ.

3.3 Localization

Recall that a multiplicatively closed subset of a ring R is a set $W \ni 1$ that is closed for products. The three most important classes of multiplicatively closed sets are the following:

Example 3.67. Let R be a commutative ring.

1) For any $f \in R$, the set $W = \{1, f, f^2, f^3, \dots\}$ is a multiplicatively closed set.

2) If $P \subseteq R$ is a prime ideal, the set $W = R \setminus P$ is multiplicatively closed: this is an immediate translation of the definition.

3) An element that is not a zerodivisor is called a **nonzerodivisor** or **regular element**. The set of regular elements in R forms a multiplicatively closed subset. When R is a domain, this set is precisely the set of all nonzero elements $R \setminus \{0\}$.

Definition 3.68 (Localization of a ring)**.** Let R be a commutative ring, and W be a multiplicative set with $0 \notin W$. The **localization** of R at W is a ring, denoted by $W^{-1}R$ or R_W, given by where $\sim$ is the equivalence relation

$$\frac{r}{w} \sim \frac{r'}{w'} \text{ if there exists } u \in W \text{ such that } u(rw' - r'w) = 0.$$

The operations are given by

$$\frac{r}{v} + \frac{s}{w} = \frac{rw + sv}{vw} \qquad \text{and} \qquad \frac{r}{v}\frac{s}{w} = \frac{rs}{vw}.$$

The zero in $W^{-1}R$ is $\frac{0}{1}$ and the multiplcative identity is $\frac{1}{1}$. There is a canonical ring homomorphism

$$\begin{aligned} R &\longrightarrow W^{-1}R\,. \\ r &\longmapsto \tfrac{r}{1} \end{aligned}$$

Note that we write elements in $W^{-1}R$ in the form $\frac{r}{w}$ even though they are equivalence classes of such expressions.

Let M be an R-module. The **localization** of M at W is the $W^{-1}R$-module $W^{-1}M$ or M_W given by

$$W^{-1}M := \left\{ \frac{m}{w} \;\middle|\; m \in M, w \in W \right\} / \sim$$

where $\sim$ is the equivalence relation $\dfrac{m}{w} \sim \dfrac{m'}{w'}$ if $u(mw' - m'w) = 0$ for some $u \in W$. The operations are given by

$$\frac{m}{v} + \frac{n}{w} = \frac{mw + nv}{vw} \qquad \text{and} \qquad \frac{r}{v}\frac{m}{w} = \frac{rm}{vw}.$$

The zero in the module $W^{-1}M$ is given by $\frac{0}{1}$.

Here are the most important examples of localizations you will come across in commutative algebra.

Example 3.69 (Most important localizations)**.** Let R be a commutative ring.

1) For $f \in R$ and $W = \{1, f, f^2, f^3, \dots\} = \{f^n \mid n \geqslant 0\}$, we usually write R_f for $W^{-1}R$.

2) When W is the set of nonzerodivisors on R, we call $W^{-1}R$ the **total ring of fractions** of R. When R is a domain, this is just the fraction field of R, and in this case this coincides with the localization at the prime (0), as described below.

3) For a prime ideal P in R, we generally write R_P for $(R \setminus P)^{-1}R$, and call it **the localization of R at P**. Given an ideal I in R, we sometimes write I_P to refer to IR_P, the image of I via the canonical map $R \to R_P$. Notice that when we localize at a prime P, the resulting ring is a local ring (R_P, P_P). We can think of the process of localization at P as *zooming in* at the prime P. Many properties of an ideal I can be checked *locally*, by checking them for IR_P for each prime $P \in V(I)$.

Remark 3.70. If R is a domain, the equivalence relation defining the localization simplifies to $rw' = r'w$. In particular, $\operatorname{Frac}(R) = R_{(0)} = (R \setminus \{0\})^{-1}R$ is a localization of R.

If R is not a domain, the canonical map $R \to W^{-1}R$ is not necessarily injective.

Example 3.71. Consider $R = k[x, y]/(xy)$. The canonical maps $R \longrightarrow R_{(x)}$ and $R \longrightarrow R_y$ are not injective, since in both cases y is invertible in the localization, and thus

$$x \mapsto \frac{x}{1} = \frac{xy}{y} = \frac{0}{y} = \frac{0}{1}.$$

In $W^{-1}R$, every element of W becomes a unit. The following universal property says roughly that $W^{-1}R$ is the smallest R-algebra in which every element of W is a unit.

Theorem 3.72. *Let R be a commutative ring, and W a multiplicative set with $0 \notin W$. Let S be an R-algebra in which every element of W is a unit. Then there is a unique homomorphism α such that the following diagram commutes:*

$$\begin{array}{ccc} R & \longrightarrow & W^{-1}R \\ \downarrow & \swarrow_{\alpha} & \\ S & & \end{array}$$

where the vertical map is the structure homomorphism and the horizontal map is the canonical homomorphism.

Proof. Given an R-algebra S such that every element of W is a unit, where the algebra structure is induced by the ring homomorphism $f\colon R \to S$, consider the map

$$\begin{aligned} W^{-1}R &\xrightarrow{\quad\alpha\quad} S \\ \tfrac{r}{w} &\longmapsto f(w)^{-1}f(r). \end{aligned}$$

First, note that our assumption that every element of W is invertible in S means that $f(w)$ is invertible in S, and thus $f(w)^{-1}f(r)$ makes sense. Moreover, we claim that α is a ring homomorphism:

$$\alpha(1) = f(1)^{-1}f(1) = 1,$$

and moreover

$$\begin{aligned}\alpha\left(\frac{a}{u}\frac{b}{v}\right)\alpha\left(\frac{ab}{uv}\right)&\\ &= f(uv)^{-1}f(ab)\\ &= (f(u)^{-1}f(a))(f(v)^{-1}f(b))\\ &= \alpha\left(\frac{a}{u}\right)\left(\frac{b}{v}\right)\end{aligned}$$

and

$$\begin{aligned}\alpha\left(\frac{a}{u}+\frac{b}{v}\right)\alpha\left(\frac{av+bu}{uv}\right)&\\ &= f(uv)^{-1}f(av+bu)\\ &= (f(u)^{-1}f(v)^{-1})(f(a)f(v)+f(b)f(u))\\ &= (f(u)^{-1}f(a)+(f(v)^{-1}f(b)\\ &= \alpha\left(\frac{a}{u}\right)+\left(\frac{b}{v}\right).\end{aligned}$$

Our definition of α gives us

$$\alpha\left(\frac{r}{1}\right) = f(1)^{-1}f(r) = f(r),$$

as desired. Moreover, if $\beta\colon W^{-1}R \to S$ is any ring homomorphism such that

$$\beta(\frac{r}{1}) = f(1)^{-1}f(r) = f(r),$$

then

$$\beta\left(\frac{r}{s}\right) = \beta\left(\frac{s}{1}\right)^{-1}\beta\left(\frac{r}{1}\right) = f(s)^{-1}f(r) = \alpha\left(\frac{s}{1}\right)^{-1}\alpha\left(\frac{r}{1}\right) = \alpha\left(\frac{r}{s}\right).$$

This proves our uniqueness claim. □

Definition 3.73. Let R be a commutative ring and let W be a multiplicative subset of R. The **localization at** W is the functor R-**Mod** $\to W^{-1}R$-**Mod** that sends each R-module M to the $W^{-1}R$-module $W^{-1}M$, and that sends each R-module homomorphism $f\colon M \to N$ to the homomorphism of $W^{-1}R$-modules given by

$$\begin{aligned}W^{-1}M &\longrightarrow W^{-1}N\\ \tfrac{m}{w} &\longmapsto \tfrac{f(m)}{w}.\end{aligned}$$

We might denote this functor by $W^{-1}(-)$ or $(-)_W$. When W is the complement of a prime ideal P, we write the localization *at* P as $(-)_P$.

Exercise 3.74. Show that for all R-module homomorphisms $f\colon M \to N$,

$$\begin{aligned}W^{-1}M &\longrightarrow W^{-1}N\\ \tfrac{m}{w} &\longmapsto \tfrac{f(m)}{w}\end{aligned}$$

is a homomorphism of modules over $W^{-1}R$.

Exercise 3.75. Show that localization is an exact additive functor.

Theorem 3.76. *Let R be a commutative ring, and $W \ni 1$ a multiplicative subset of R. Then the localization at W and $W^{-1}R \otimes -$ are naturally isomorphic functors. In particular, for every R-module M, there is an isomorphism of $W^{-1}R$-modules*

$$W^{-1}R \otimes_R M \cong W^{-1}M,$$

and given an R-module map $\alpha\colon M \to N$, the map of $W^{-1}R$-modules $W^{-1}R \otimes \alpha$ corresponds to $W^{-1}\alpha = \alpha_W$ under these isomorphisms.

Proof. The bilinear map
$$\begin{aligned} W^{-1}R \times M &\longrightarrow W^{-1}M \\ (\tfrac{r}{w}, m) &\longmapsto \tfrac{rm}{w} \end{aligned}$$
induces a homomorphism $\psi\colon W^{-1}R \times M \to W^{-1}M$ that is surjective.

For an inverse map, set $\phi(\frac{m}{w}) := \frac{1}{w} \otimes m$. To see this is well-defined, suppose $\frac{m}{w} = \frac{m'}{w'}$, so there exists some $v \in W$ such that $v(mw' - m'w) = 0$. Then,

$$\phi\left(\frac{m}{w}\right) - \phi\left(\frac{m'}{w'}\right) = \frac{1}{w} \otimes m - \frac{1}{w'} \otimes m'.$$

We can multiply through by $\frac{vww'}{vww'}$ to get

$$\frac{vw'}{vww'} \otimes m - \frac{vw}{vww'} \otimes m' = \frac{1}{vww'} \otimes v(mw' - m'w) = 0.$$

To see this is a homomorphism, we note that

$$\begin{aligned} \phi\left(\frac{m}{w} + \frac{m'}{w'}\right) = \phi\left(\frac{mw' + m'w}{ww'}\right) &= \frac{1}{ww'} \otimes (mw' + m'w) = \frac{1}{ww'} \otimes mw' + \frac{1}{ww'} \otimes m'w \\ &= \frac{w'}{ww'} \otimes m + \frac{w}{ww'} \otimes m' = \frac{1}{w} \otimes m + \frac{1}{w'} \otimes m' = \phi\left(\frac{m}{w}\right) + \phi\left(\frac{m'}{w'}\right), \end{aligned}$$

and

$$\phi\left(r\frac{m}{w}\right) = \frac{1}{w} \otimes rm = r\left(\frac{1}{w} \otimes m\right) = r\phi\left(\frac{m}{w}\right).$$

The composition $\phi \circ \psi$ sends

$$\frac{r}{w} \otimes m \mapsto \frac{rm}{w} \mapsto \frac{1}{w} \otimes rm = \frac{r}{w} \otimes m.$$

Since this is the identity on simple tensors, and simple tensors generated the tensor product, it must be the identity.

For the claim about maps, we need check that $\psi_N \circ (W^{-1}R \otimes \alpha) = W^{-1}\alpha \circ \psi_M$ for every R-module homomorphism $\alpha! : M \to N$. And indeed,

$$\begin{aligned} (\psi_N \circ (W^{-1}R \otimes \alpha))\left(\frac{r}{w} \otimes m\right) = \psi_N\left(\frac{r}{w} \otimes \alpha(m)\right) &= \frac{r\alpha(m)}{w} \\ &= \frac{\alpha(rm)}{w} = W^{-1}\alpha\left(\frac{rm}{w}\right) = (W^{-1}\alpha \circ \psi_M)\left(\frac{r}{w} \otimes m\right). \end{aligned}$$

Finally, we note that our isomorphisms $W^{-1}R \otimes_R M \cong W^{-1}M$ give a natural isomorphism between the localization functor $W^{-1}(-)$ and the tensor functor $W^{-1}R \otimes_R -$. Indeed, given a map of R-modules $M \xrightarrow{f} N$, the diagram

$$\begin{array}{ccc} W^{-1}R \otimes M & \xrightarrow{\psi_M} & W^{-1}M \\ \downarrow{\scriptstyle \mathrm{id} \otimes f} & & \downarrow{\scriptstyle W^{-1}(f)} \\ W^{-1}R \otimes N & \xrightarrow[\psi_N]{} & W^{-1}N \end{array}$$

commutes, since it commutes for simple tensors:

$$\begin{array}{ccc} \frac{r}{w} \otimes m & \longrightarrow & \frac{rm}{w} \\ \downarrow{\scriptstyle \mathrm{id} \otimes f} & & \downarrow{\scriptstyle W^{-1}(f)} \\ \frac{r}{w} \otimes f(m) & \longrightarrow & \frac{rf(m)}{w} = \frac{f(rm)}{w}. \end{array}$$

□

Now since localization is exact, we conclude that $W^{-1}R \otimes_R -$ is an exact functor for all commutative rings R and all multiplicatively closed subsets W.

Exercise 3.77. Let R be a commutative noetherian ring, W be a multiplicative set, M be a finitely generated R-module, and N an arbitrary R-module. Show that

$$\mathrm{Hom}_{W^{-1}R}(W^{-1}M, W^{-1}N) \cong W^{-1}\,\mathrm{Hom}_R(M, N).$$

In particular, if P is prime,

$$\mathrm{Hom}_{R_P}(M_P, N_P) \cong \mathrm{Hom}_R(M, N)_P.$$

Localization is a very powerful tool in commutative algebra. Many important concepts localize well, in the sense that to prove that R or a module satisfy a certain property, it is often sufficient to show that all localizations of R or of that module also have that property. This is a very common and helpful technique in commutative algebra. For example, a module M is zero if and only if all its localizations are zero; one can even reduce to showing all localizations of M at a prime ideal are zero.

One important thing to keep in mind, however, is that if M is a finitely generated R-module, a localization M_W of M is typically not finitely generated over R, though it is finitely generated over R_W.

Exercise 3.78. Let R be a domain and let $f \in R$ be a nonzero nonunit. Then R_f is not a finitely generated R-module.

To solve this exercise, however, one needs a little bit of commutative algebra that we are not covering in this course.

3.4 Hom-tensor adjunction

The Hom and tensor functors are closely related. First, we note that $\mathrm{Hom}_R(A, B)$ can be a module over a ring S when A or B have a bimodule structure.

Exercise 3.79. Let R and S be rings.

- If A is an (R, S)-bimodule and B is a left R-module, then $\mathrm{Hom}_R(A, B)$ has a left S-module structure via $(s \cdot f)(a) = f(as)$.
- If A is an (R, S)-bimodule and B is a right S-module, then $\mathrm{Hom}_S(A, B)$ has a right R-module structure via $(f \cdot r)(a) = f(ra)$.
- If B is an (S, R)-bimodule and A is a right R-module, then $\mathrm{Hom}_R(A, B)$ has a left S-module structure via $(s \cdot f)(a) = sf(a)$.
- If B is an (S, R)-bimodule and A is a left S-module, then $\mathrm{Hom}_S(A, B)$ has a right R-module structure via $(f \cdot r)(a) = f(a)r$.

These structures can be a bit confusing at first – especially since we have left module structures written on the right and vice-versa. While the exercise is not difficult, it can be extremely enlightening – we strongly recommend the reader tries their hand at the details.

The following statements are known as Hom-tensor adjunction – and as we will see, they do encode an adjunction of functors.

Theorem 3.80. *Let R and S be rings. Assume that*

- *A is a right R-module,*
- *B is an (R, S)-bimodule, and*
- *C is a right S-module.*

There is a natural isomorphism of abelian groups

$$\mathrm{Hom}_S(A \otimes_R B, C) \cong \mathrm{Hom}_R(A, \mathrm{Hom}_S(B, C)).$$

If A also has a (T, R)-bimodule structure, or C has a (T, S)-bimodule structure, then this is an isomorphism of (left or right, respectively) T-modules.

Theorem 3.81. *Let R and S be rings. Assume that*

- *A is a left R-module,*
- *B is an (S, R)-bimodule, and*
- *C is a left S-module.*

There is a natural isomorphism of abelian groups

$$\mathrm{Hom}_S(B \otimes_R A, C) \cong \mathrm{Hom}_R(A, \mathrm{Hom}_S(B, C)).$$

We leave the details to the reader, and prove the case when the underlying rings are commutative. First, let's do the case when $R = S$.

Theorem 3.82 (Hom-tensor adjunction I)**.** *Let R be a commutative ring and let M, N, and P be R-modules. There is an isomorphism of R-modules*

$$\mathrm{Hom}_R(M \otimes_R N, P) \cong \mathrm{Hom}_R(M, \mathrm{Hom}_R(N, P))$$

that is natural on M, N, and P.

Proof. The universal property of the tensor product says that to give an R-module homomorphism $M \otimes_R N \longrightarrow P$ is the same as giving an R-bilinear map $M \times N \longrightarrow P$. Given such a bilinear map f, the map $n \mapsto f(m \otimes n)$ is R-linear for each $m \in M$, so it defines an R-module homomorphism $N \longrightarrow P$. Now the assignment

$$\begin{aligned} M &\longrightarrow \mathrm{Hom}_R(N, P) \\ m &\longrightarrow (n \mapsto f(m \otimes n)) \end{aligned}$$

is R-linear, f is an R-module homomorphism, and $m \mapsto m \otimes n$ is R-linear on m.

Conversely, given an R-module homomorphism $f \in \mathrm{Hom}_R(M, \mathrm{Hom}_R(N, P))$, one can check (exercise!) that $(m, n) \mapsto f(m)(n)$ is an R-bilinear map, so it induces an R-module homomorphism $M \otimes_R N \longrightarrow P$. Moreover, the two constructions are inverse to each other.

So we have constructed a bijection of Hom-sets

$$\begin{aligned} \mathrm{Hom}_R(M \otimes_R N, P) &\xrightarrow{\ \tau\ } \mathrm{Hom}_R(M, \mathrm{Hom}_R(N, P)) \,. \\ f &\longmapsto (m \mapsto (n \mapsto f(m \otimes n))) \\ (m \otimes n \mapsto g(m)(n)) &\longmapsfrom g \end{aligned}$$

It's routine to check that both of these bijections are indeed homomorphisms of R-modules, so we leave it as an exercise.

Finally, naturality means we have the following commutative diagrams:

$$\begin{array}{ccc} A & & \mathrm{Hom}_R(A \otimes_R N, P) \xrightarrow{\cong} \mathrm{Hom}_R(A, \mathrm{Hom}_R(N, P)) \,, \\ {\scriptstyle f}\downarrow & \rightsquigarrow & {\scriptstyle (f \otimes 1_N)^*}\uparrow \qquad\qquad\qquad {\scriptstyle f^*}\uparrow \\ B & & \mathrm{Hom}_R(B \otimes_R N, P) \xrightarrow[\cong]{} \mathrm{Hom}_R(B, \mathrm{Hom}_R(N, P)) \end{array}$$

$$\begin{array}{ccc} A & & \mathrm{Hom}_R(M \otimes_R A, P) \xrightarrow{\cong} \mathrm{Hom}_R(M, \mathrm{Hom}_R(A, P)) \,, \\ {\scriptstyle f}\downarrow & \rightsquigarrow & \uparrow{\scriptstyle (1_M \otimes f)^*} \qquad\qquad\qquad \uparrow{\scriptstyle (f^*)_*} \\ B & & \mathrm{Hom}_R(M \otimes_R B, P) \xrightarrow[\cong]{} \mathrm{Hom}_R(M, \mathrm{Hom}_R(B, P)) \end{array}$$

and

$$\begin{array}{ccc} A & & \mathrm{Hom}_R(M \otimes_R N, A) \xrightarrow{\cong} \mathrm{Hom}_R(M, \mathrm{Hom}_R(N, A)) \,. \\ {\scriptstyle f}\downarrow & \rightsquigarrow & {\scriptstyle f_*}\downarrow \qquad\qquad\qquad \downarrow{\scriptstyle (f^*)^*} \\ B & & \mathrm{Hom}_R(M \otimes_R N, B) \xrightarrow[\cong]{} \mathrm{Hom}_R(M, \mathrm{Hom}_R(N, B)) \end{array}$$

We leave checking these do indeed commute as an exercise. □

Corollary 3.83 (Tensor and Hom are adjoint functors)**.** *Let R be a commutative ring, and M an R-module. The functor $- \otimes_R M \colon R\text{-}\boldsymbol{Mod} \longrightarrow R\text{-}\boldsymbol{Mod}$ is left adjoint to the functor $\mathrm{Hom}_R(M, -) \colon R\text{-}\boldsymbol{Mod} \longrightarrow R\text{-}\boldsymbol{Mod}$.*

Proof. The adjointness translates into the fact that for all R-modules N and P there is a bijection

$$\mathrm{Hom}_R(N \otimes_R M, P) \cong \mathrm{Hom}_R(N, \mathrm{Hom}_R(M, P))$$

which is natural on N and P, which is a corollary of Theorem 3.82. □

Later, when we talk about more general abelian categories, we will see that this adjunction *implies* that Hom is left exact and that tensor is right exact; in fact, this is a more general fact about adjoint pairs. For now, we want to discuss a more general version of this Hom-tensor adjunction.

Theorem 3.84 (Hom-tensor adjunction II)**.** *Let $f : R \to S$ be a ring homomorphism of commutative rings. Let M be an R-module, and P and N be S-modules. There is an isomorphism of abelian groups*

$$\mathrm{Hom}_S(M \otimes_R N, P) \cong \mathrm{Hom}_R(M, \mathrm{Hom}_S(N, P)).$$

Moreover, this isomorphism is natural on M, N, and P, so it induces natural isomorphisms

- *between* $\mathrm{Hom}_S(- \otimes_R N, P)$ *and* $\mathrm{Hom}_R(-, \mathrm{Hom}_S(N, P))$.
- *between* $\mathrm{Hom}_S(M \otimes_R -, P)$ *and* $\mathrm{Hom}_R(M, \mathrm{Hom}_S(-, P))$.
- *between* $\mathrm{Hom}_S(M \otimes_R N, -)$ *and* $\mathrm{Hom}_R(M, \mathrm{Hom}_S(N, -))$.

Proof. Consider the map

$$\begin{aligned} \mathrm{Hom}_S(M \otimes_R N, P) &\xrightarrow{\tau} \mathrm{Hom}_R(M, \mathrm{Hom}_S(N, P)) \,. \\ f &\longmapsto m \mapsto (n \mapsto f(m \otimes n)) \end{aligned}$$

Fix f. For each $m \in M$, let τ_m be the map $N \longrightarrow P$ defined by $\tau_m(n) := f(m \otimes n)$. Note that τ_m is indeed a homomorphism of S-modules, since it is the composition of two S-module maps, f and $m \otimes_R \mathrm{id}_N$, where m is the constant map $M \longrightarrow M$ equal to m.

We should check that our proposed map τ is indeed a map of abelian groups. It is immediate from the definition that τ sends the 0-map to the 0-map. Moreover, given S-module homomorphisms $f, g \colon M \otimes N \longrightarrow P$, and any $n \in N$, we have

$$\begin{aligned} \tau_m(f+g)(n) &= (f+g)(m \otimes n) && \text{by definition} \\ &= f(m \otimes n) + g(m \otimes n) && \text{since } f \text{ and } g \text{ are } S\text{-module maps} \\ &= \tau_m(f)(n) + \tau_m(g)(n) && \text{by definition} \end{aligned}$$

so $\tau_m(f+g) = \tau_m(f) + \tau_m(g)$ for all $m \in M$, and thus $\tau(f+g) = \tau(f) + \tau(g)$.

Suppose that $\tau(f) = 0$. Then for every $m \in M$ and every $n \in N$,

$$0 = \tau(f)(m)(n) = \tau_m(f)(n) = f(m \otimes n),$$

so f vanishes at every simple tensor, and we must have $f = 0$. On the other hand, if we are given $g \in \mathrm{Hom}_R(M, \mathrm{Hom}_S(N, P))$, consider the map $M \times N \longrightarrow P$ defined by $\tilde{f}(m, n) = g(m)(n)$. Since g is a homomorphism of R-modules, it is R-linear on m. Moreover, for each fixed m, $g(m)$ is a homomorphism of S-modules, so in particular $g(m)$ is R-linear. Together, these say that $\tilde{f}$ is an R-bilinear map. Let f be the homomorphism of R-modules $M \otimes_R N \longrightarrow P$ induced by $\tilde{f}$. By definition, $f(m \otimes n) = \tilde{f}(m, n) = g(m)(n)$, so $\tau(f) = g$. We conclude that τ is a bijection.

We leave the statements about naturality as exercises. □

Corollary 3.85 (Adjointness of restriction and extension of scalars)**.** *Let $f\ R \longrightarrow S$ be a ring homomorphism. The restriction of scalars functor $f^*\colon S$-**Mod** $\longrightarrow R$-**Mod** is the right adjoint of the extension of scalars functor $f_*\colon R$-**Mod** $\longrightarrow S$-**Mod**.*

Proof. We need to show that for every R-module M and every S-module N there are bijections

$$\mathrm{Hom}_S(f_*(M), N) \cong \mathrm{Hom}_R(M, f^*(N))$$

which are natural on both M and N. By Theorem 3.84, we have natural bijections

$$\mathrm{Hom}_S(M \otimes_R S, N) \cong \mathrm{Hom}_R(M, \mathrm{Hom}_S(S, N)).$$

The module $M \otimes_R S$ is precisely $f_*(M)$. By Theorem 3.4, $\mathrm{Hom}_S(S, N) \cong N$ as an S-module. An isomorphism of S-modules $\mathrm{Hom}_S(S, N) \longrightarrow N$ is in particular an R-linear map, and thus also an isomorphism of R-modules. So $\mathrm{Hom}_S(S, N) \cong f^*(N)$ as R-modules. Therefore, the Hom-tensor adjuntion gives us the natural bijections we were looking for. □

The idea is that restriction of scalars and extension of scalars are the most efficient ways of making an R-module out of an S-module, and vice-versa.

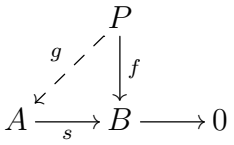
Chapter 4

Enough (about) projectives and injectives

While Hom and tensor are not exact functors in general, $\mathrm{Hom}_R(M, -)$, $\mathrm{Hom}_R(-, M)$, and $M \otimes_R -$ can be exact functors for carefully chosen modules M. In this chapter, we introduce these three classes of modules (projective, injective, and flat modules) and study their properties. Throughout, we consider general rings and left modules.

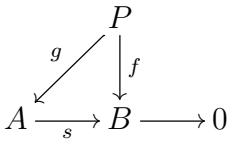
4.1 Projectives

Definition 4.1. Let R be a ring. An R-module P is **projective** if given any surjective R-module homomorphism $s\colon A \to B$ and any R-module homomorphism $f\colon P \to B$, there exists an R-module homomorphism g such that the diagram



commutes.

Remark 4.2. The commutativity of the diagram



says that $s_*(g) = f$, where s_* is the map $\mathrm{Hom}_R(P, A) \longrightarrow \mathrm{Hom}_R(P, B)$ induced by s. Whenever this happens, we say that g is a **lifting** of f, and that f **lifts**, or that f **factors through** A.

There are projective modules over any ring, as the next result shows; in fact, free modules are always projective.

Theorem 4.3. *Free modules are projective.*

Proof. Let F be a free R-module. Suppose we are given R-module homomorphisms $s\colon A \to C$ and $f\colon F \to C$ such that s is surjective. Fix a basis $B = \{b_i\}_i$ for F. Since s is surjective, for each i we can choose $a_i \in A$ such that $s(a_i) = f(b_i)$. Consider the function $u\colon B \longrightarrow A$ given by $u(b_i) = a_i$. The universal property of free modules says that there exists an R-module homomorphism $g\colon F \longrightarrow A$ that coincides with u for all basis elements. Now

$$sg(b_i) = su(b_i) = s(a_i) = f(b_i),$$

so sg agrees with f for all basis elements. Since B generates F, we conclude that $sg = f$. □

Projective modules are precisely those that make the covariant Hom functor exact.

Theorem 4.4. *Let P be an R-module. The functor $\mathrm{Hom}_R(P, -)$ is exact if and only if P is projective.*

Proof. By Theorem 3.26, $\mathrm{Hom}_R(P, -)$ is left exact. The statement is that P is projective if and only for any short exact sequence

$$0 \longrightarrow A \xrightarrow{\ i\ } B \xrightarrow{\ p\ } C \longrightarrow 0\ ,$$

the induced map $s_*\colon \mathrm{Hom}_R(P, B) \longrightarrow \mathrm{Hom}_R(P, C)$ is surjective. Say we are given a surjective map

$$B \xrightarrow{\ p\ } C \longrightarrow 0\ .$$

The induced map s_* is surjective if and only if for every $f \in \mathrm{Hom}_R(P, C)$ there exists a lifting $g \in \mathrm{Hom}_R(P, B)$ of f, meaning $s_*(g) = f$. By Theorem 4.2, the existence of such a g for all such surjective maps s is precisely the condition that P is projective. □

Corollary 4.5. *For any ring R, $\mathrm{Hom}_R(R, -)$ is exact. More generally, if F is any free R-module, then $\mathrm{Hom}_R(F, -)$ is exact.*

Proof. By Theorem 4.3, free modules, and R in particular, are projective. By Theorem 4.4, $\mathrm{Hom}_R(F, -)$ must be exact for any free R-module F. □

However, not every projective module is free. But before we see such examples, we need to know a bit more about projective modules.

First, we show that we can rephrase the condition that a module is projective or injective in terms of split exact sequences.

Theorem 4.6. *An R-module P is projective if and only if every short exact sequence*

$$0 \longrightarrow A \longrightarrow B \longrightarrow P \longrightarrow 0$$

splits.

Proof. ($\Rightarrow$) Consider a short exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} P \longrightarrow 0\,.$$

If P is projective, the identity map on P lifts to a map $P \longrightarrow B$, meaning that

$$\begin{array}{ccccc} & & & & P \\ & & & \swarrow^{h} & \| \\ B & \xrightarrow[g]{} & P & \longrightarrow & 0 \end{array}$$

commutes. This says that our map h

$$0 \longrightarrow A \xrightarrow{f} B \underset{h}{\overset{g}{\rightleftarrows}} P \longrightarrow 0$$

is a splitting for our short exact sequence, which must then be split, by Theorem 2.26.

($\Leftarrow$) Conversely, suppose that every short exact sequence

$$0 \longrightarrow A \longrightarrow B \longrightarrow P \longrightarrow 0$$

splits, and consider any diagram

$$\begin{array}{ccccc} & & & & P \\ & & & & \downarrow{\scriptstyle f} \\ B & \xrightarrow[p]{} & C & \longrightarrow & 0. \end{array}$$

Let F be a free module that surjects onto P — for example, the free module on a set of generators of P — and fix a surjection $\pi\colon F \twoheadrightarrow P$. By assumption, the short exact sequence

$$0 \longrightarrow \ker \pi \longrightarrow F \underset{h}{\overset{\pi}{\rightleftarrows}} P \longrightarrow 0$$

splits, so by Theorem 2.26 there exists h such that $\pi h = \mathrm{id}_P$. Now since F is free, we can define an R-module map $\hat{g}\colon F \longrightarrow B$ that such that

$$\begin{array}{ccccc} & & \overset{h}{\curvearrowleft} & & \\ F & \xrightarrow[\pi]{} & P & & \\ {\scriptstyle \hat{g}}\downarrow & \swarrow^{g} & \downarrow{\scriptstyle f} & & \\ B & \xrightarrow[s]{} & C & \longrightarrow & 0 \end{array}$$

commutes, by sending each basis element $b \in F$ to any lift of $f\pi(b)$ in B via s. Now set $g := \hat{g}h$, and note that

$$\begin{aligned} sg &= s\hat{g}h && \text{by definition} \\ &= f\pi h && \text{by commutativity} \\ &= f && \text{since } \pi h = \mathrm{id}_P, \end{aligned}$$

so g is a lift of s by f. □

We have seen that free modules are projective; what other modules are projective?

Definition 4.7. An R-module M is a **direct summand** of an R-module N if there exists an R-module A such that $A \oplus M \cong N$.

Remark 4.8. Saying that M is a direct summand of N is equivalent to giving a splitting π of the inclusion map $i\colon M \hookrightarrow N$, meaning that $\pi i = \mathrm{id}_N$. As we have argued in Theorem 2.26, such a splitting π gives

$$N = \operatorname{im} i \oplus \ker \pi.$$

Essentially repeating the argument we used in Theorem 2.26, every element in N can be written as

$$n = (n - i\pi(n)) + i\pi(n),$$

where $i\pi(n) \in \operatorname{im} i$ and $n - i\pi(n) \in \ker \pi$, and $\ker \pi \cap \operatorname{im} i = 0$ because if $i(a) \in \ker \pi$ then $a = \pi i(a) = 0$.

Note that when we are dealing with graded modules over a graded ring, the kernels and images of graded maps are graded modules, and the equality $N = \operatorname{im} i \oplus \ker \pi$ is a graded direct sum of graded modules.

Theorem 4.9. *An R-module is projective P if and only if P is a direct summand of a free R-module. In particular, a finitely generated R-module P is projective if and only if P is a direct summand of R^n for some n.*

Proof. ($\Rightarrow$) Let P be a projective module, and fix a free module F surjecting onto P. If P is finitely generated, we can take $F = R^n$ for some n. The short exact sequence

$$0 \longrightarrow \ker \pi \longrightarrow F \xrightarrow{\pi} P \longrightarrow 0$$

must split by Theorem 4.6, so P is a direct summand of F.

($\Leftarrow$) Now suppose P is a direct summand of a free module F. In particular, we have an inclusion map $i\colon P \longrightarrow F$ that splits, so it comes together with a projection map $\pi\colon F \longrightarrow P$ such that $\pi i = \mathrm{id}_P$. Given any diagram

$$\begin{array}{ccccc} & & & & P \\ & & & & \downarrow{\scriptstyle f} \\ B & \xrightarrow[s]{} & C & \longrightarrow & 0, \end{array}$$

we can define an R-module homomorphism h such that $sh = f\pi$, so that the following diagram commutes:

$$\begin{array}{ccccc} & & \overset{i}{\curvearrowleft} & & \\ F & \xrightarrow[\pi]{} & P & & \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle f} & & \\ B & \xrightarrow[s]{} & C & \longrightarrow & 0, \end{array}$$

Setting $g := hi$, we do indeed obtain $sg = f$, since

$$\begin{aligned} sg &= shi && \text{by definition} \\ &= f\pi i && \text{because } sh = f\pi \\ &= f && \text{since } \pi i = \mathrm{id}_P . \end{aligned}$$

□

Remark 4.10. While every module M is a quotient of a free module F, so that we always have a surjection $\pi\colon F \to M$, not every module embeds into a free module; and even if M is a submodule of some free module F, the inclusion map $M \subseteq F$ is not necessarily split. On the other hand, as we showed in Theorem 4.9 that M is projective if and only if we can write it as a quotient of a free module F, say $\pi\colon F \to M$, and π splits, so that in fact M embeds into F and that map splits.

Corollary 4.11. *Let R be any ring.*

1) Every direct summand of a projective module is projective.

2) Every direct sum of projective modules is projective.

Proof.

1) Suppose $M \oplus A \cong P$ for some projective module P. By Theorem 4.9, there exists a free R-module F and an R-module B such that $P \oplus B \cong F$. Then $M \oplus A \oplus B \cong P \oplus B \cong F$, and by Theorem 4.9 this implies M is projective.

2) Let $\{P_i\}_{i \in I}$ be a family of projective modules. By Theorem 4.9, there exist free modules F_i such that each P_i is a direct summand of F_i. Therefore, $\oplus P_i$ is a direct summand of $\oplus_i F_i$, which is also free. By Theorem 4.9, this implies that $\oplus P_i$ is projective. □

We are finally ready to give examples showing that projective modules are not necessarily free.

Example 4.12. The ring $R = \mathbb{Z}/(6)$ can be written as a direct sum of the ideals

$$I = (2) \text{ and } J = (3).$$

Indeed, $R = I + J$ and $I \cap J = 0$, so $R = I \oplus J$. By Theorem 4.11, I and J are projective R-modules. However, I and J are *not* free. This can easily be explained numerically: every finitely generated free R-module is of the form R^n, so it has 6^n elements for some n, while I and J have 3 and 2 elements respectively.

Finally, to emphasize its importance we record here an easy result that we have used repeatedly at this point, and which will be very important later on.

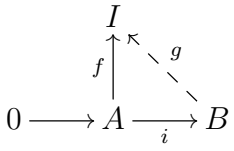
Lemma 4.13. *For every R-module M, there exists a free module F surjecting onto M. If M is finitely generated, we can take F to be finitely generated.*

We will often need only a weaker version of this: that every module is a quotient of a *projective* module.

4.2 Injectives

Injective modules are dual to projectives.

Definition 4.14. An R-module I is **injective** if given an injective R-module homomorphism $i : A \longrightarrow B$ and an R-module homomorphism $f : A \longrightarrow I$, there exist an R-module homomorphism g such that



commutes.

These are precisely the modules I such that $\mathrm{Hom}_R(-, I)$ is exact.

Theorem 4.15. *An R-module I is injective if and only if $\mathrm{Hom}_R(-, I)$ is exact, meaning that for every short exact sequence*

$$0 \longrightarrow A \xrightarrow{i} B \xrightarrow{p} C \longrightarrow 0$$

we get an exact sequence

$$0 \longrightarrow \mathrm{Hom}_R(C, I) \xrightarrow{p^*} \mathrm{Hom}_R(B, I) \xrightarrow{i^*} \mathrm{Hom}_R(A, I) \longrightarrow 0.$$

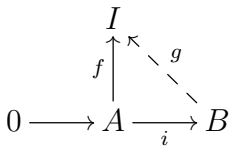
Proof. By Theorem 3.26, $\mathrm{Hom}_R(-, I)$ is left exact, so for any short exact sequence

$$0 \longrightarrow A \xrightarrow{i} B \xrightarrow{p} C \longrightarrow 0$$

we get an exact sequence

$$0 \longrightarrow \mathrm{Hom}_R(C, I) \xrightarrow{p^*} \mathrm{Hom}_R(B, I) \xrightarrow{i^*} \mathrm{Hom}_R(A, I).$$

So the content of the theorem is that I is injective if and only if for every injective R-module homomorphism $i : A \longrightarrow B$, the induced map i^* is surjective. Now notice that i^* is surjective if and only if every $f \in \mathrm{Hom}_R(A, I)$ lifts to some $g \in \mathrm{Hom}_R(B, I)$, meaning



commutes. That is precisely what we want for I to be injective. □

Giving examples of injective modules is much harder than giving examples of projective modules, but we will see some examples later. First, we prove some properties of injective modules.

The class of injectives modules is closed for products and finite direct sums.

Lemma 4.16. *Given any family $\{M_i\}_{i\in I}$ of injective modules, $\prod_{i\in I} M_i$ is injective.*

Proof. Let $\pi_j \colon \prod_{i\in I} M_i \longrightarrow M_j$ be the projection onto the jth factor. Given any diagram

$$\begin{array}{ccccc} & & \prod_{i\in I} M_i & & \\ & & \uparrow{\scriptstyle f} & & \\ 0 & \longrightarrow & A & \xrightarrow[i]{} & B, \end{array}$$

the fact that M_i is injective gives us R-module homomorphisms g_i such that

$$\begin{array}{ccccc} & & M_i & & \\ & & \uparrow{\scriptstyle \pi_i f} \quad \nwarrow{\scriptstyle g_i} & & \\ 0 & \longrightarrow & A & \xrightarrow[i]{} & B \end{array}$$

commutes for each i. Now the R-module homomorphism

$$\begin{aligned} B &\xrightarrow{g} \textstyle\prod_{i\in I} M_i \\ b &\longmapsto (g_i(b)) \end{aligned}$$

makes the diagram

$$\begin{array}{ccccc} & & \prod_{i\in I} M_i & & \\ & & \uparrow{\scriptstyle f} \quad \nwarrow{\scriptstyle g} & & \\ 0 & \longrightarrow & A & \xrightarrow[i]{} & B \end{array}$$

commute, so $\prod_{i\in I} M_i$ is injective. □

Lemma 4.17. *If $M \oplus N = E$ is an injective R-module, then so are M and N.*

Proof. Any diagram

$$\begin{array}{ccccc} & & M & & \\ & & \uparrow{\scriptstyle f} & & \\ 0 & \longrightarrow & A & \xrightarrow[i]{} & B \end{array}$$

can be extended to a map $A \longrightarrow E$ by composing f with the inclusion of the first factor. Since E is injective, there exists h such that

$$\begin{array}{ccccc} & & M & \xrightarrow{j} & E \\ & & \uparrow{\scriptstyle f} & & \uparrow{\scriptstyle h} \\ 0 & \longrightarrow & A & \xrightarrow[i]{} & B \end{array}$$

commutes. Let $\pi : E \longrightarrow M$ be the projection onto M, so that $\pi j = \mathrm{id}_M$. Now if we set $g := \pi h$,

$$\begin{array}{ccccc} & & M & \xrightarrow{\pi} & E \\ & & \uparrow{\scriptstyle f} & & \uparrow{\scriptstyle h} \\ 0 & \longrightarrow & A & \xrightarrow[i]{} & B \end{array} \qquad \begin{aligned} gi &= \pi h i && \text{by definition} \\ &= \pi j f && \text{by commutativity} \\ &= f && \text{because } \pi j = \mathrm{id}_M . \end{aligned}$$

□

Theorem 4.18 (Baer Criterion)**.** *An R-module E is injective if and only if every R-module homomorphism $I \longrightarrow E$ from an ideal I in R can be extended to the whole ring, meaning that there exists g making the diagram*

$$\begin{array}{ccccc} & & E & & \\ & & \uparrow{\scriptstyle f} & \nwarrow^{g} & \\ 0 & \longrightarrow & I & \longrightarrow & R \end{array}$$

commute.

Proof. On the one hand, if E is injective then our condition is simply a special case of the definition of injective module. On the other hand, suppose that this condition holds, and consider any diagram

$$\begin{array}{ccccc} & & E & & \\ & & \uparrow{\scriptstyle f} & & \\ 0 & \longrightarrow & M & \longrightarrow & N. \end{array}$$

To simplify notation, let's assume our map $M \longrightarrow N$ is indeed the inclusion of the submodule M, so we can write $m \in N$ for the image of m in N. Consider the set

$$X := \{(A, g) \mid A \text{ is a submodule of } N, M \subseteq A \subseteq N, \text{ and } g \text{ extends } f\}.$$

First, notice X is nonempty, since $(M, f) \in X$. Moreover, we can partially order X by setting $(A, g) \leqslant (B, h)$ if $A \subseteq B$ and $h|_A = g$. So we have a nonempty partially ordered set; let's show we can apply Zorn's Lemma to it.

Given a chain in X, meaning a sequence

$$(A_1, g_1) \leqslant (A_2, g_2) \leqslant \cdots$$

of nested submodules $A_1 \subseteq A_2 \subseteq \cdots$ and maps g_i that extend all g_j with $j \leqslant i$, let $A := \bigcup_i A_i$, and define

$$\begin{aligned} A &\xrightarrow{\quad g \quad} E \\ a &\longrightarrow g_i(a) \text{ if } a \in A_i. \end{aligned}$$

Since all the g_i are homomorphisms of R-modules, this map g is indeed a map of R-modules. Moreover, g is well-defined, since the $g_i(a) = g_j(a)$ whenever $a \in A_i \cap A_j$. By construction, this map extends all the g_i, so we conclude that (A, g) is an upper bound for our chain. Moreover, $M \subseteq A \subseteq N$ follows immediately from our construction, and since each g_i extends f, so does g. We conclude that $(A, g) \in X$, and more generally that any chain in X has an upper bound in X. So Zorn's Lemma applies.

By Zorn's Lemma, X has a maximal element, say (A, g). We claim that $A = N$. Suppose not, and let $n \in N$ be an element not in A. One can easily check that

$$I := \{r \in R \mid rn \in A\}$$

is an ideal in R, and that

$$\begin{aligned} I &\xrightarrow{\quad h \quad} E \\ r &\longrightarrow g(rn) \end{aligned}$$

is an R-module homomorphism.

By assumption, we can extend h to an R-module homomorphism $R \longrightarrow E$, which we will write as h as well. Now the R-module homomorphism

$$\begin{aligned} A + Rn &\xrightarrow{\varphi} E \\ a + rn &\longrightarrow g(a) + h(r) \end{aligned}$$

is well-defined by construction, since any $rn \in A$ satisfies $g(rn) = h(r)$, and if $rn = r'n$ then $h(r) = rn = r'n = h(r')$. Finally, this map agrees with g on A, and thus it agrees with f on M, so $(A + Rn, \varphi) \in X$ and $(A, g) \leqslant (A + Rn, \varphi)$. By the maximality of (A, g), we conclude that $A + Rn = A$, and thus $n \in A$, which is a contradiction. We conclude that $A = N$. Therefore, g makes the diagram

$$\begin{array}{ccccc} & & E & & \\ & & \uparrow{\scriptstyle f} & \nwarrow^{g} & \\ 0 & \longrightarrow & M & \longrightarrow & N. \end{array}$$

commute. □

Over a noetherian ring, an arbitrary sum of injective modules is still injective.

Corollary 4.19. *Let R be a noetherian ring. If $\{M_j\}_{j \in J}$ are all injective R-modules, then so is $\bigoplus_{j \in J} M_j$.*

Proof. By Theorem 4.18, it is enough to show that any R-module map

$$f \colon I \to \bigoplus_{j \in J} M_j$$

from an ideal I extends to R. Since R is noetherian, I is finitely generated, so let $I = (a_1, \ldots, a_n)$. For each $i = 1, \ldots, n$, the element $f(a_i) = (b_{i,j})_{j \in J}$ has $b_{i,j} \neq 0$ only for finitely many values of $j \in J$. Then

$$K := \{j \in J \mid f(a_i)_j \neq 0 \text{ for some } i = 1, \ldots, n\}$$

is a finite set, and $f(I) \subseteq \bigoplus_{j \in K} M_j$. Direct sums of finitely many modules coincide with their product, so by Theorem 4.16, $\bigoplus_{j \in K} M_j$ is injective. Therefore, there exists g such that

$$\begin{array}{ccccc} & & \bigoplus_{k \in K} M_k & & \\ & & \uparrow{\scriptstyle f} & \nwarrow^{g} & \\ 0 & \longrightarrow & I & \longrightarrow & R \end{array}$$

commutes. Now $\bigoplus_{k \in K} M_k$ is a submodule of $\bigoplus_{j \in J} M_j$, so we can think of g as an R-module homomorphism with codomain $\bigoplus_{j \in J} M_j$, and

$$\begin{array}{ccccc} & & \bigoplus_{j \in J} M_j & & \\ & & \uparrow{\scriptstyle f} & \nwarrow^{g} & \\ 0 & \longrightarrow & I & \longrightarrow & R \end{array}$$

commutes. □

We can now give some interesting examples of injective modules.

Theorem 4.20. *If R is a domain, then the fraction field $Q = R_{(0)}$ is an injective R-module.*

Proof. By Baer's Criterion, we just need to show that every R-module homomorphism $f : I \to Q$ can be extended to R. If $I = 0$, f is the zero map, so we can simply extend it to the zero map $I \to Q$. So suppose that $I \neq 0$. First, note that for any nonzero $a, b \in I$,

$$af(b) = f(ab) = bf(a),$$

so

$$\frac{1}{a}f(a) = \frac{1}{b}f(b).$$

So let

$$c := \frac{1}{a}f(a) \quad \text{for any nonzero } a \in I.$$

Let

$$\begin{array}{ccc} R & \xrightarrow{g} & Q \\ r & \longmapsto & rc. \end{array}$$

In other words, g is the R-module homomorphism determined by setting $1 \mapsto c$. Now we claim that g extends f: indeed, given any $a \in I$,

$$g(a) = ac = a\frac{1}{a}f(a) = f(a).$$

Thus by Baer's Criterion, Q must be an injective R-module. □

By Theorem 4.17, finite direct sums of injective modules are injective; when R is noetherian, we can take arbitrary direct sums. So when R is a noetherian domain, and Q is its fraction field, any Q-vector space is also an injective R-module. In fact, one can remove the noetherianity assumption, and prove directly that every Q-vector space is an injective R-module by using an argument similar to the one in Theorem 4.20.

Exercise 4.21. Show that if R is a domain and Q is its fraction field, then every Q-vector space is an injective R-module.

It is elementary to show that every R-module is a quotient of a free module. The dual statement is true as well, but it is a lot more delicate. That's our next goal; our work starts with divisible modules.

Definition 4.22. An R-module D is **divisible** if for every nonzero $r \in R$ and every $d \in D$ there exists $b \in D$ such that $rb = d$.

Remark 4.23. Given $r \in R$, and an R-module M, the multiplication by r map $M \xrightarrow{\cdot r} M$ is an R-module homomorphism. The R-module M is divisible if and only if multiplication by r is surjective for all nonzero $r \in R$.

Lemma 4.24. *Any quotient of a divisible module is also divisible.*

Proof. Let D be a divisible R-module and E be a submodule of D. Let $r \in R$ and $d + E \in D/E$. By assumption, there exists $a \in D$ such that $ra = d$. The image $a + E$ of a in D/E is still a solution to $r(a + E) = d + D$, so indeed E is divisible. □

Lemma 4.25. *Over a domain, every injective module is divisible.*

Proof. Suppose that E is an injective R-module, where R is a domain. Fix $r \in R$ and $a \in E$. Since R is a domain, we have $sr = s'r \Rightarrow s = s'$ for any $s, s', r \in R$. In particular, each element in (r) can be written uniquely as sr for some $s \in R$. In particular, the map of R-modules

$$\begin{aligned} (r) &\longrightarrow E \\ sr &\longrightarrow sa \end{aligned}$$

is well-defined. Since E is injective, we can extend this to a homomorphism $f : R \longrightarrow E$. Finally, $f(1) \in E$ is an element such that $e = f(r) = rf(1)$, and E is divisible. □

This not true in general if we do not assume R is a domain.

Example 4.26. Let k be a field and $R = k[x]/(x^2)$. On the one hand, R is not a divisible R-module, since there is no $y \in R$ such that $xy = 1$. On the other hand, R is actually an injective module over itself, although we do not have the tools to justify that this is indeed an injective R-module.[1]

The converse of Theorem 4.25 does not hold in general, and quotients of injective modules are not necessarily injective.

Exercise 4.27. Let $R = k[x, y]$, where k is a field, let $Q = \mathrm{frac}(R)$ be the fraction field of R. The R-module $M = Q/R$ is divisible but not injective.

But the converse of Theorem 4.25 does hold for some special classes of rings.

Lemma 4.28. *Let R be a principal ideal domain. An R-module E is injective if and only E is divisible.*

Proof. Given Theorem 4.25, we only need to show that divisible modules are injective. By Baer's Criterion, we only need to show that any map from an ideal to E can be extended to the whole ring. So let E be a divisible R-module, and consider any map $I \longrightarrow E$ from an ideal I to E. If $I = 0$, we could extend our map by taking the 0 map from R to E, so we might as well assume that $I \neq 0$. By assumption, $I = (a)$ for some $a \in R$, and since E is divisible, there exists $e \in E$ such that $f(a) = ae$. Now consider the multiplication by r map,

$$\begin{aligned} R &\xrightarrow{g} E \\ r &\longrightarrow re. \end{aligned}$$

For every $r \in R$, $g(ra) = rae = rf(a) = f(ra)$, so g extends f. Therefore, by Theorem 4.18, E is injective. □

[1] Using fancy words you might learn in Commutative Algebra II, this ring R is an example of a complete intersection, which is a subclass of Gorenstein rings. Moreover, $\dim R = 0$ – this is something you'd learn about in Commutative Algebra II. Now it turns out (and this is a nontrivial fact) that Gorenstein rings of dimension 0 are injective modules over themselves.

Lemma 4.29. *Over a principal ideal domain, quotients of injective modules are injective.*

Proof. If E is injective, it is also divisible, by Theorem 4.25. Given any submodule $D \subseteq E$, any $e \in E$, and a nonzero $r \in R$, there exists $y \in E$ such that $ry = e$, and so this also holds in E/D. Then E/D is divisible, and thus injective by 4.28. □

Given an injective abelian group, we can always use it to construct an injective R-module over our favorite ring R.

Lemma 4.30. *Given an injective abelian group D and a ring R, $\mathrm{Hom}_{\mathbb{Z}}(R, D)$ is an injective R-module.*

Proof. Let $E := \mathrm{Hom}_{\mathbb{Z}}(R, D)$. This abelian group E is a left R-module, via

$$r \cdot f := (a \mapsto f(ar)).$$

We claim that E is actually an injective R-module. By Theorem 4.15, it is sufficient to prove that $\mathrm{Hom}_R(-, \mathrm{Hom}_{\mathbb{Z}}(R, D))$ is an exact functor. By Hom-tensor adjunction, the functor $\mathrm{Hom}_R(-, \mathrm{Hom}_{\mathbb{Z}}(R, D))$ is naturally isomorphic to $\mathrm{Hom}_{\mathbb{Z}}(- \otimes_R R, D)$. This last functor is a composition:

$$\mathrm{Hom}_{\mathbb{Z}}(- \otimes_{\mathbb{Z}} R, D) = \mathrm{Hom}_{\mathbb{Z}}(-, D) \circ (- \otimes_R R).$$

On the one hand, $- \otimes_R R$ is naturally isomorphic to the identity on R-**Mod**, by Theorem 3.55, so it is exact. On the other hand, D is an injective $\mathbb{Z}$-module, so $\mathrm{Hom}_{\mathbb{Z}}(-, D)$ is exact by Theorem 4.15. The composition of exact functors is exact, and thus $\mathrm{Hom}_R(-, \mathrm{Hom}_{\mathbb{Z}}(R, D))$ is exact. □

Example 4.31. Since $\mathbb{Q}$ is a divisible abelian group, by Theorem 4.30 for any ring R the R-module $\mathrm{Hom}_{\mathbb{Z}}(R, \mathbb{Q})$ is injective.

When we talked about projective modules, we showed that every module is a quotient of a projective – in fact, every module is a quotient of a free module. The dual statement is true as well: that every module embeds into an injective module. We will soon see that these two statements are extremely important.

While the statement about projectives is relatively simple – it's essentially a consequence of the universal property of free modules – the fact about injectives is a lot more delicate; the work we just did on divisible modules was precisely so we could show this deep and important fact.

First, we show that every abelian group can be embedded into an injective abelian group.

Lemma 4.32. *Every abelian group M is a submodule of some injective abelian group.*

Proof. On the one hand, M is a quotient of some free abelian group, say $M \cong (\oplus_i \mathbb{Z})/K$. Now $\mathbb{Z}$ embeds in $\mathbb{Q}$, and thus M embeds into a quotient of $\oplus_i \mathbb{Q}$. By Theorem 4.20, $\mathbb{Q}$ is an injective abelian group, and by Theorem 4.19, $\oplus_i \mathbb{Q}$ is an injective abelian group, since $\mathbb{Z}$ is a noetherian ring. By Theorem 4.29, any quotient of $\oplus_i \mathbb{Q}$ is also injective, so we have shown that M embeds into an injective abelian group, say D. □

In fact, the proof above can be repeated over any PID: if R is a PID, we can show that any R-module M embeds into an injective module, and in fact M embeds into some number of copies of the fraction field Q.

We can finally show that over any ring, every module can be embedded into an injective module.

Theorem 4.33. *Every R-module M is a submodule of some injective R-module.*

Proof. First, by Theorem 4.32 we can view M as a subgroup of some injective abelian group D. Let $i\colon M \longrightarrow D$ be the inclusion map and $E := \mathrm{Hom}_{\mathbb{Z}}(R, D)$.

By Theorem 4.30, E is an injective R-module. Since Hom is left exact, by Theorem 3.26, $\mathrm{Hom}_{\mathbb{Z}}(R, -)$ preserves the inclusion I, so we have an inclusion $\mathrm{Hom}_{\mathbb{Z}}(R, M) \subseteq \mathrm{Hom}_{\mathbb{Z}}(R, D)$. Now consider the map

$$\begin{aligned} M &\xrightarrow{\psi} \mathrm{Hom}_{\mathbb{Z}}(R, M) \\ m &\longrightarrow (r \mapsto rm). \end{aligned}$$

This is an R-module homomorphism:

- Given $a, b \in M$,
$$\psi(a+b)(r) = r(a+b) = ra + rb = \psi(a)(r) + \psi(b)(r),$$
so $\psi(a+b) = \psi(a) + \psi(b)$.
- Given $r \in R$, $m \in M$, and $s \in R$,
$$\psi(rm)(s) = s(rm) = r(sm) = r\psi(m)(s),$$
so $\psi(rm) = r\psi(m)$.

Moreover, if $\psi(m) = 0$ then $m = \psi(m)(1) = 0$. So ψ is injective, and thus composing ψ with our previous inclusion $\mathrm{Hom}_{\mathbb{Z}}(R, M) \subseteq \mathrm{Hom}_{\mathbb{Z}}(R, D)$ gives us an inclusion φ of M into the injective R-module $\mathrm{Hom}_{\mathbb{Z}}(R, D)$. However, the inclusion $\mathrm{Hom}_{\mathbb{Z}}(R, M) \subseteq \mathrm{Hom}_{\mathbb{Z}}(R, D)$ is a priori only a map of abelian groups, so we should check that φ is indeed R-linear. In order to do this, we need to be careful (at least in the case when R is not commutative) with how we defined the left R-module structure on $\mathrm{Hom}_{\mathbb{Z}}(R, D)$ in Theorem 3.79: this is a situation where we view R as a $(\mathbb{Z}, R)$-bimodule and D as a left $\mathbb{Z}$-module, so $\mathrm{Hom}_{\mathbb{Z}}(R, D)$ is a left R-module via

$$r \cdot f \text{ is the } R\text{-map given by } (r \cdot f)(a) = f(ar).$$

The map we need to show is R-linear is

$$\begin{aligned} M &\xrightarrow{\varphi} \mathrm{Hom}_{\mathbb{Z}}(R, D) \\ m &\longmapsto \varphi_m = (r \mapsto i(rm)). \end{aligned}$$

Regarding i as a simple inclusion, $i(m)$ simply views the element m as an element of D; to simplify notation, we drop the i: so for each $m \in M$, $\varphi(m)$ is the map $\varphi_m\colon R \longrightarrow D$ given by

$$\varphi_m(r) = rm.$$

For every $r \in R$, $m \in M$, and $s \in R$,

$$\begin{aligned} \varphi_{rm}(s) &= s(rm) && \text{by definition} \\ &= (sr)m && \text{using the module axioms} \\ &= \varphi_m(sr) && \text{by definition} \\ &= r\varphi_m(s) && \text{by definition of the left } R\text{-module structure on } \mathrm{Hom}_{\mathbb{Z}}(R, D) \end{aligned}$$

so $\varphi(rm) = r\varphi(m)$. This shows that φ is an inclusion of R-modules. □

And finally, just like we did for projectives, we can characterize injectives in terms of split short exact sequences.

Theorem 4.34. *An R-module I is injective if and only if every short exact sequence*

$$0 \longrightarrow I \longrightarrow B \longrightarrow C \longrightarrow 0$$

splits.

Proof. Let I be an injective R-module, and consider any short exact sequence

$$0 \longrightarrow I \xrightarrow{i} B \xrightarrow{p} C \longrightarrow 0.$$

Since I is injective, there exists a map g making

$$\begin{array}{ccccc} & & I & & \\ & & \| & \nwarrow^{g} & \\ 0 & \longrightarrow & I & \xrightarrow[i]{} & B \end{array}$$

commute, and such a g gives a splitting for our short exact sequence.

Conversely, suppose that every short exact sequence $0 \longrightarrow I \longrightarrow B \longrightarrow C \longrightarrow 0$ splits, and consider a diagram

$$\begin{array}{ccccc} & & I & & \\ & & {\scriptstyle f}\uparrow & & \\ 0 & \longrightarrow & A & \xrightarrow[i]{} & B. \end{array}$$

By Theorem 4.33, I embeds into some injective R-module E, say by the inclusion j. By assumption, the short exact sequence

$$0 \longrightarrow I \xrightarrow{j} E \longrightarrow \operatorname{coker} j \longrightarrow 0$$

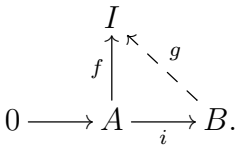
splits, so there exists a map $q\colon E \longrightarrow I$ such that $qi = \mathrm{id}_I$. Since E is injective, we can lift i through jf, obtaining an R-module homomorphism ℓ such that

$$\begin{array}{ccccc} & & I & \xrightarrow[j]{\overset{q}{\curvearrowleft}} & E \\ & & {\scriptstyle f}\uparrow & & \uparrow{\scriptstyle \ell} \\ 0 & \longrightarrow & A & \xrightarrow[i]{} & B \end{array}$$

commutes. Now $g := q\ell$ satisfies

$$\begin{aligned} gi &= q\ell i && \text{by definition} \\ &= qjf && \text{by commutativity} \\ &= f && \text{since } qj = \mathrm{id}_I, \end{aligned}$$

so



commutes. □

Before we move on from injective modules, let us say a word about how the story continues. The next chapter is quite beautiful, and it is a shame we have no time to discuss it in detail this semester.

We proved above that every module M is a submodule of some injective module. One can even do better and talk about the smallest injective module that M embeds in; this is called the **injective hull** $E(M)$ of M. One could describe $E(M)$ by saying that it is the intersection of all the injective modules that contain M, but this is not a very practical description. Injective hulls can also be described through the theory of essential extensions, a topic which we do not have time to discuss this semester. We leave the definition here just for fun, but we do not have the time to talk about it at length.

Definition 4.35. Let $M \subseteq E$. We say E is an *essential extension* of M if every nonzero submodule $N \subseteq E$ intersects M nontrivially, meaning $E \cap M \neq 0$. More generally, an injective map $\alpha\colon M \longrightarrow E$ is an essential extension if $\alpha(M) \subseteq E$ is an essential extension in the sense above.

One then shows that an R-module M is injective if and only if it has no proper essential extensions $E \supseteq M$. This proves that a maximal essential extension E of M is injective, and that there are no other injective modules I with of $M \subseteq I \subseteq E$. Moreover, one can show that any two maximal essential extension of M are isomorphic – and thus we can talk about the maximal essential extension of M, up to isomorphism, which is

But the theory of injectives, and injective hulls in particular, is much more complicated than the theory of projectives. When M is a finitely generated module, we can always find a *finitely generated* projective (even free!) module surjecting onto M; in contrast, the injective hull $E(M)$ might not be finitely generated – in fact, $E(M)$ is typically not finitely generated even when M is cyclic.

The story of the structure of injective modules then continues in a beautiful way. Over a noetherian ring, it turns out that every injective module can be decomposed into a direct sum of injective modules of the form $E(R/P)$, where P is a prime ideal in R. Moreover, the injective modules $E(R/P)$ are the indecomposable injective modules, so the basic building blocks of injective modules. One can in fact compute the injective hull of any finitely generated R-module very explicitly. A lot of this was proved in Eben Matlis' beautiful PhD thesis [Mat58], but sadly we do not have time for the details this semester. The details, however, are very important, for example to develop the theory of local cohomology – a topic which we will briefly mention later on.

4.3 Flat modules

Finally, we turn to the modules that make the tensor product exact.

Definition 4.36. An R-module M is said to be **flat** if $M \otimes_R -$ is an exact functor.

Remark 4.37. By Theorem 3.64, $M \otimes_R -$ is right exact. Therefore, M is flat if and only if for every injective R-module map $i\colon A \longrightarrow B$,

$$M \otimes_R A \xrightarrow{1\otimes i} M \otimes_R B \quad \text{is injective.}$$

Lemma 4.38. *Given a family of R-modules $\{M_i\}_{i\in I}$, the direct sum $\oplus_i M_i$ is flat if and only if every M_i is flat. In particular, direct summands of flat modules are flat.*

Proof. Given a family of R-module homomorphisms $f_i\colon M_i \longrightarrow N_i$, there is an R-module homomorphism

$$\begin{aligned} \textstyle\bigoplus_{i\in I} M_i &\xrightarrow{(f_i)_{i\in I}} \bigoplus_{i\in I} N_i \\ (m_i) &\longmapsto (f_i(m_i)) \end{aligned}$$

which is injective if and only if every f_i is injective.

Let $f : A \longrightarrow B$ be an injective R-module homomorphism. There is a commutative diagram

$$\begin{array}{ccc} \left(\bigoplus_{i\in I} M_i\right) \otimes_R A & \xrightarrow{\cong} & \bigoplus_{i\in I} M_i \otimes_R A \\ \downarrow{\scriptstyle \varphi := 1\otimes f} & & \downarrow{\scriptstyle (1\otimes f)_i =: \psi} \\ \left(\bigoplus_{i\in I} M_i\right) \otimes_R B & \xrightarrow[\cong]{} & \bigoplus_{i\in I} M_i \otimes_R B \end{array}$$

where the horizontal maps are the isomorphisms from Theorem 3.56. In particular, φ is injective if and only if ψ is injective. Moreover, ψ is injective if and only if each component is injective, meaning $1 \otimes f\colon M_i \otimes A \longrightarrow M_i \otimes B$ is injective for all i.

On the one hand, $\bigoplus_{i\in I} M_i$ is flat if and only if for every injective map f, the corresponding ϕ is injective. On the other hand, all the M_i are flat if and only if for every injective map f, $1 \otimes f\colon M_i \otimes A \longrightarrow M_i \otimes B$ is injective for all i, or equivalently, as explained above, if ψ is injective for any given injective map f. This translates into the equivalence we want to show. □

All projectives are flat.

Theorem 4.39. *Let R be any ring. Every projective R-module is flat.*

Proof. First, recall that $R \otimes_R -$ is naturally isomorphic to the identity functor, by Theorem 3.55, and thus exact (see Theorem 3.22). This shows that R is flat, and thus any free module, being a direct sum of copies of R, must also be flat by Theorem 4.38. Finally, every projective module is a direct summand of a free module, by Theorem 4.9. Direct summands of flat modules are flat, by Theorem 4.38, so every projective module is flat. □

We can test whether a given module if flat by looking at the finitely generated submodules.

Theorem 4.40. *If every finitely generated submodule of M is flat, then M is flat.*

Proof. Let $i\colon A \longrightarrow B$ be an injective map of R-modules. We want to show that

$$M \otimes_R A \xrightarrow{1\otimes i} M \otimes_R B$$

is injective. Suppose that $u \in \ker(1_M \otimes i)$. We are going to construct a finitely generated submodule $N \subseteq M$, with $j\colon N \to M$ the inclusion, and an element $v \in N \otimes_R A$ such that $v \in \ker(1_N \otimes i)$ and $u = (j \otimes 1_A)(v)$. Once we do that, our submodule N is finitely generated, and thus flat by assumption, so $1_N \otimes i$ is injective; therefore, $v = 0$ and thus we must have $u = 0$. Therefore, $1_M \otimes i$ is injective, and we conclude that M is flat.

Let's say that $u = m_1 \otimes a_1 + \cdots + m_n \otimes a_n$. In Theorem 3.34, we constructed the tensor product $M \otimes_R B$ as a quotient of the free abelian group F on $M \times B$ by the subgroup S with all the necessary relations we need to impose. This gives us a short exact sequence

$$0 \longrightarrow S \longrightarrow F \xrightarrow{\pi} M \otimes_R B \longrightarrow 0.$$

The fact that $m_1 \otimes i(a_1) + \cdots + m_n \otimes i(a_n) = 0$ means we can rewrite this element as $\pi(s)$ for some $s \in S$. This element s is can be written as a finite sum $s = q_1 + \cdots + q_r$ of elements of the form

$$(m, b + b') - (m, b) - (m, b'), (m + m', b) - (m, b) - (m', b), \text{ or } (mr, b) - (m, rb).$$

Now let C be the set obtained by collecting the M-coordinates of elements of $M \times B$ appearing in any of the q_i, as follows:

- if $q_i = (m, b + b') - (m, b) - (m, b')$, we let $m \in C$;
- if $q_i = (m + m', b) - (m, b) - (m', b)$, we let $m, m' \in C$;
- and if $q_i = (mr, b) - (m, rb)$.

For each i, we pick only one such representation of q_i, so that C is a finite set. Let $c_1, \ldots, c_t$ be the finitely many elements of C, and take N to be the finitely generated submodule of M generated by $m_1, \ldots, m_n$ and $c_1, \ldots, c_t$. Note that $q_1 + \cdots + q_r$ is an element of the free module on $N \times B$, and thus

$$m_1 \otimes i(a_1) + \cdots + m_n \otimes i(a_n) = 0 \quad \text{in } N \otimes_R B.$$

Consider $v = m_1 \otimes a_1 + \cdots + m_n \otimes a_n \in N \otimes_R A$. Now

$$(j \otimes 1_A)(v) = (j \otimes 1_A)(m_1 \otimes a_1 + \cdots + m_n \otimes a_n) = m_1 \otimes a_1 + \cdots + m_n \otimes a_n \in M \otimes_R A,$$

and

$$(1_N \otimes i)(v) = (1_N \otimes i)(m_1 \otimes a_1 + \cdots + m_n \otimes a_n) = m_1 \otimes i(a_1) + \cdots + m_n \otimes i(a_n) = 0,$$

as desired.

The reason we needed to add in these extra elements is that a priori $N \otimes_R B$ is not necessarily a submodule of $M \otimes B$, so we do not necessarily have $m_1 \otimes i(a_1) + \cdots + m_n \otimes i(a_n) = 0$ in $(Rm_1 + \cdots + Rm_n) \otimes B$ without adding in all relations that make it true. □

Definition 4.41. Let R be a domain and M be an R-module. The **torsion submodule** of M is

$$T(M) := \{m \in M \mid rm = 0 \text{ for some regular element } r \in R\}.$$

The elements of $T(M)$ are called **torsion elements**, and we say that M **is torsion** if $T(M) = M$. Finally, M is **torsion free** if $T(M) = 0$.

Lemma 4.42. *If R is a domain and M is a flat R-module, then M is torsion free.*

Proof. Let $Q = \text{frac}(R)$ be the fraction field of R, which is a torsion free R-module. Now $M \otimes_R Q$ is a Q-vector space, so isomorphic to a direct sum of copies of Q. In particular, $M \otimes_R Q$ is torsion free as an R-module. Since M is flat, the inclusion $R \subseteq Q$ induces an injective R-module map

$$0 \longrightarrow M \otimes_R R \longrightarrow M \otimes_R Q,$$

and since $M \cong M \otimes_R R$, by Theorem 3.55, we conclude that M is isomorphic to a submodule of $M \otimes_R Q$. Submodules of torsion free modules are also torsion free, so M is torsion free. □

In general, the converse does not hold.

Example 4.43. Let k be a field and $R = k[x, y]$. Consider the ideal $\mathfrak{m} = (x, y)$. This is a submodule of the torsion free module R, and thus $\mathfrak{m}$ is torsion free. However, it is not flat. Suppose, by contradiction, that $\mathfrak{m}$ is a flat module; then when we apply $\mathfrak{m} \otimes_R -$ to the inclusion $\mathfrak{m} \subseteq R$ we must get an inclusion

$$\mathfrak{m} \otimes_R \mathfrak{m} \xrightarrow{\varphi} \mathfrak{m} \otimes R \cong \mathfrak{m}\,.$$

However, the element $x \otimes y - y \otimes x \in \mathfrak{m} \otimes_R \mathfrak{m}$ is taken to

$$\varphi(x \otimes y - y \otimes x) = x \otimes y - y \otimes x = xy \otimes 1 - yx \otimes 1 = 0,$$

while we claim that $x \otimes y - y \otimes x$ is not zero in $\mathfrak{m} \otimes_R \mathfrak{m}$.

This can be done in the usual way, by setting up an R-bilinear map $\mathfrak{m} \times \mathfrak{m} \to R/\mathfrak{m}$, and showing that the resulting homomorphism of R-modules $\mathfrak{m} \otimes_R \mathfrak{m} \to R/\mathfrak{m}$ sends $x \otimes y - y \otimes x$ to something nonzero. Since this point of our study of homological algebra we'd like to avoid boring arguments about R-biadditive maps, here is an alternative proof.

Consider the canonical quotient map $\mathfrak{m} \to \mathfrak{m}/\mathfrak{m}^2$, and note that $\mathfrak{m}/\mathfrak{m}^2$ can also be viewed as a 2-dimensional $R/\mathfrak{m}$-vector space with generators $e_1 := x + \mathfrak{m}$ and $e_2 := y + \mathfrak{m}$. Since tensor is right exact, by Theorem 3.64, we get an induced surjection

$$\mathfrak{m} \otimes_R \mathfrak{m} \twoheadrightarrow \mathfrak{m}/\mathfrak{m}^2 \otimes_R \mathfrak{m}.$$

Similarly, tensoring with $\mathfrak{m}/\mathfrak{m}^2$ is also right exact, so we get a surjection

$$\mathfrak{m}/\mathfrak{m}^2 \otimes_R \mathfrak{m} \twoheadrightarrow \mathfrak{m}/\mathfrak{m}^2 \otimes_R \mathfrak{m}/\mathfrak{m}^2.$$

Composing the two, we get a surjection

$$\mathfrak{m} \otimes_R \mathfrak{m} \twoheadrightarrow \mathfrak{m}/\mathfrak{m}^2 \otimes_R \mathfrak{m}/\mathfrak{m}^2$$

which sends $x \otimes y - y \otimes x$ to $e_1 \otimes e_2 - e_2 \otimes e_1$. By Theorem 4.44,

$$\mathfrak{m}/\mathfrak{m}^2 \otimes_R \mathfrak{m}/\mathfrak{m}^2 \cong \mathfrak{m}/\mathfrak{m}^2 \otimes_{R/\mathfrak{m}} \mathfrak{m}/\mathfrak{m}^2;$$

the isomorphism sends $e_1 \otimes e_2 - e_2 \otimes e_1$ to $e_1 \otimes e_2 - e_2 \otimes e_1$. By Theorem 3.58, $e_1 \otimes e_2$ and $e_2 \otimes e_1$ are linearly independent, and thus $e_1 \otimes e_2 - e_2 \otimes e_1 \neq 0$. We conclude that $x \otimes y - y \otimes x \neq 0$ in $\mathfrak{m} \otimes_R \mathfrak{m}$.

Above we used the following elementary but useful fact about tensor products:

Exercise 4.44. Let R be a commutative ring, I an ideal in R, and let M and N be R-modules such that $I \subseteq \operatorname{ann}(M) \cap \operatorname{ann}(N)$. Show that there is an isomorphism

$$M \otimes_R N \cong M \otimes_{R/I} N.$$

The converse of Theorem 4.42 does hold over a PID.

Lemma 4.45. *If R is a principal ideal domain, an R-module M is flat if and only if it is torsion free.*

Proof. The fact that flat implies torsion free is a special case of Theorem 4.42. So suppose M is a torsion free R-module. First, we will deal with the case when M is finitely generated. The structure theorem for PIDs says that M must be isomorphic to a direct sum of copies of cyclic modules. The cyclic module R/I has torsion – all the elements are killed by I – unless $I = 0$. Therefore, M must be isomorphic to a direct sum of copies of R, and thus free. By Theorem 4.3, M is projective, and by Theorem 4.39 projectives are flat, so M is flat.

Now let M be any torsion free R-module. All of the finitely generated submodules of R are also torsion free, and thus flat by what we have shown above. By Theorem 4.40, M must be flat. □

But not all flat modules are projective.

Example 4.46. The $\mathbb{Z}$-module $\mathbb{Q}$ is torsion free and thus flat, by Theorem 4.45. However, $\mathbb{Q}$ is not a projective $\mathbb{Z}$-module. Suppose, by contradiction, that $\mathbb{Q}$ is a projective $\mathbb{Z}$-module. By Theorem 4.9, $\mathbb{Q}$ must be a direct summand of a free module, say $F = \bigoplus_I \mathbb{Z}$. Consider the inclusion $\iota : \mathbb{Q} \hookrightarrow F$, and pick $i \in I$ such that the image of $\mathbb{Q}$ contains some element with a nonzero entry in the i component. Now consider the projection $\pi : F \longrightarrow \mathbb{Z}$ onto the ith factor. By assumption, the composition $\pi i : \mathbb{Q} \longrightarrow \mathbb{Z}$ is nonzero. However, there are no nontrivial abelian group homomorphisms $\mathbb{Q} \longrightarrow \mathbb{Z}$, contradicting the fact that πi is nonzero. We conclude that $\mathbb{Q}$ is not projective.

For finitely generated modules over a commutative noetherian local ring, every flat module is free, and thus flat, projective, and free all coincide. However, to prove that we need a little bit of commutative algebra, which we introduce in the next section.

Theorem 4.47 (Flatness of localization)**.** *Let R be a commutative ring, and $W \ni 1$ a multiplicative subset of R. Then $W^{-1}R$ is flat over R.*

Proof. By Theorem 3.76, tensoring with $W^{-1}R$ is localizing at W. But localization is exact, so tensoring with $W^{-1}R$ is exact, and thus $W^{-1}R$ is a flat R-module. □

So for example, if R is a domain then its fraction field Q is a flat module.

Definition 4.48. An R-module F is **faithfully flat** if F is flat and $F \otimes_R M \neq 0$ for every nonzero R-module M.

Exercise 4.49. Let R be a commutative ring. Show that the following are equivalent:

a) F is faithfully flat.

b) F is flat and for every maximal ideal $\mathfrak{m}$, $\mathfrak{m}F \neq F$.

c) The complex

$$A \xrightarrow{f} B \xrightarrow{g} C$$

is exact if and only if

$$F \otimes_R A \xrightarrow{1\otimes f} F \otimes_R B \xrightarrow{1\otimes g} F \otimes_R C$$

is exact.

4.4 Commutative local rings

We have shown that

$$\text{Free} \implies \text{projective} \implies \text{flat}.$$

Over a local ring, these three notions actually coincide. To show this, we need a little bit of commutative algebra. First, some notation: when R is a local ring, meaning R has a unique maximal ideal $\mathfrak{m}$, we write $(R, \mathfrak{m})$ to denote the ring R and its maximal ideal. Now note that for any R-module M, the module $M/\mathfrak{m}M$ is annihilated by $\mathfrak{m}$, so it is also a module over a ring $R/\mathfrak{m}$, which is a field.

The following is a classical result in commutative algebra, known by some as Nakayama's Lemma. As noted in [Mat89, page 8], Nakayama himself claimed that this should be attributed to Krull and Azumaya, but it's not clear which of the three actually had the commutative ring statement first. So some authors (eg, Matsumura) prefer to refer to it as NAK. There are actually a range of statements, rather than just one, that go under the banner of Nakayama's Lemma a.k.a. NAK.

Theorem 4.50 (NAK)**.** *Let $(R, \mathfrak{m}, k)$ be a local ring, and M be a finitely generated module. If $M = \mathfrak{m}M$, then $M = 0$.*

The theorem above is the theorem most commonly referred to as NAK. The proof involves only elementary tools, and a fun linear algebra-inspired trick called the Determinantal Trick. While we will not include the details here, they can be found in any standard Commutative Algebra book. We will however use this result to prove another statement that is also commonly referred to as NAK, which allows us to talk about minimal generating sets for finitely generated modules over local rings.

Remark 4.51. Let R be any commutative ring, and consider an R-module M and an ideal I. If $IM = 0$, meaning that $am = 0$ for all $a \in I$ and all $m \in M$, then M can be given the structure of an R/I-module, as follows: for any $m \in M$ and any $r \in R$,

$$(r + I)m = rm.$$

The fact that I kills M is what makes this action well-defined. The fact that M is actually an R-module under this action is a consequence of the fact that M is an R-module; checking these details is routine, and we leave them as an exercise.

Notice that the structure of M as an R/I-module is essentially *the same* as its structure as an R-module. There are many properties of M as an R-module that pass onto its R/I-module structure, and typically such results are easy to check.

Here is a special case of this: if $(R, \mathfrak{m})$ is a commutative local ring, and M is an R-module, then the module $M/\mathfrak{m}M$ is killed by $\mathfrak{m}$, and thus it is also a module over $R/\mathfrak{m}$. Now notice that $R/\mathfrak{m}$ is a field, so $M/\mathfrak{m}M$ is actually a vector space over the field $R/\mathfrak{m}$.

Theorem 4.52. *Let $(R, \mathfrak{m})$ be a commutative local ring, and M be a finitely generated module. For $m_1, \ldots, m_s \in M$,*

$$m_1, \ldots, m_s \text{ generate } M \iff \overline{m_1}, \ldots, \overline{m_s} \text{ generate } M/\mathfrak{m}M.$$

Thus, any generating set for M consists of at least $\dim_k(M/\mathfrak{m}M)$ elements.

Proof. The implication ($\Rightarrow$) is clear. For ($\Leftarrow$), given $m_1, \ldots, m_s \in M$ such that $\overline{m_1}, \ldots, \overline{m_s}$ generate $M/\mathfrak{m}M$, consider

$$N := Rm_1 + \cdots + Rm_s \subseteq M.$$

Since $M/\mathfrak{m}M$ is generated by the image of N, we have $M = N + \mathfrak{m}M$. By taking the quotient by N, we see that

$$M/N = (N + \mathfrak{m}M)/N = \mathfrak{m}\,(M/N)\,.$$

By Theorem 4.50, $M/N = 0$ and thus $M = N$. □

As we mentioned above, this allows us to talk about minimal generating sets.

Definition 4.53. Let $(R, \mathfrak{m})$ be a local ring, and M a finitely generated module. A set of elements $\{m_1, \ldots, m_t\}$ is a **minimal generating set** of M if the images of $m_1, \ldots, m_t$ form a basis for the $R/\mathfrak{m}$ vector space $M/\mathfrak{m}M$.

Note that every finitely generated module over a local ring has a minimal generating set, that every minimal generating set has the same number of elements, and that any set of generators for M contains a minimal generating set, all thanks to plain old linear algebra. In particular, we can now define the following:

Definition 4.54. Let M be a finitely generated module over a commutative local ring $(R, \mathfrak{m})$. The **minimal number of generators** of M, denoted $\mu(M)$, is the number of elements in any minimal generating set for M.

We now have the key commutative algebra ingredients needed to show that for finitely generated modules over a noetherian local ring, projective = free. However, we need one more homological tool we haven't developed yet, so we will hold off on proving this for now – in fact, you will soon be able to prove it easily, so the following problem will be in the next problem set:

Exercise 4.55. Let $(R, \mathfrak{m})$ be a commutative local ring, and let M be a finitely presented module. Then

$$M \text{ is flat} \iff M \text{ is projective} \iff M \text{ is free.}$$

Kaplansky [Kap58] showed that this holds even for modules that are not necessarily finitely presented, but generated by countably many elements.

Definition 4.56. An R-module M is **locally free** if M_P is a free R_P-module for every prime ideal P.

Exercise 4.57. Let R be a commutative ring, M and N be R-modules, and P be a prime ideal. Show that

$$(M \otimes_R N)_P \cong M_P \otimes_{R_P} N_P.$$

Exercise 4.58. Let R be a commutative ring, P be a prime ideal, and M be an R_P-module. Let N be M as an R-module via restriction of scalars. Then as R_P-modules, we have an isomorphism

$$N_P \cong M.$$

Exercise 4.59. Let R be a commutative ring. Show that a homomorphism of R-modules $f\colon M \to N$ is surjective if and only if f_P is surjective for all primes P.

Exercise 4.60. Let R be a noetherian ring, W be a multiplicative set, M be a finitely generated R-module, and N an arbitrary R-module. Show that

$$\operatorname{Hom}_{W^{-1}R}(W^{-1}M, W^{-1}N) \cong W^{-1}\operatorname{Hom}_R(M, N).$$

In particular, if P is prime,

$$\operatorname{Hom}_{R_P}(M_P, N_P) \cong \operatorname{Hom}_R(M, N)_P.$$

Theorem 4.61. *Let R be a commutative noetherian ring and let M be a finitely presented R-module. Then*

$$M \textit{ is projective} \iff M \textit{ is flat} \iff M \textit{ is locally free.}$$

Proof. We already know that projectives are flat, by Theorem 4.39.

Suppose M is flat. We claim that M_P is flat for every prime ideal P. First, note that $M_P \cong R_P \otimes_R M$, by Theorem 3.76; moreover, R_P is a flat R-module by Theorem 3.75. Note moreover that any R_P-module can also be viewed as an R-module by extension of scalars along the canonical localization map. Now given any short exact sequence of R_P-modules, say

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0,$$

tensoring with M_P over R_P can be done in two steps: first we view this as a short exact sequence of R-modules, and tensor with M, but M is a flat R-module, so

$$0 \longrightarrow A \otimes_R M \longrightarrow B \otimes_R M \longrightarrow C \otimes_R M \longrightarrow 0$$

is exact. Then we tensor with R_P, but this is also flat R-module, so we get a short exact sequence again:

$$0 \longrightarrow (A \otimes_R M) \otimes_R R_P \longrightarrow (B \otimes_R M) \otimes_R R_P \longrightarrow (C \otimes_R M) \otimes_R R_P \longrightarrow 0.$$

By Theorem 3.76 and Theorem 4.57, for each R_P-module X we have

$$(X \otimes_R M) \otimes_R R_P \cong (X \otimes_R M)_P \cong X_P \otimes_{R_P} M_P.$$

But $X_P \cong X$, by Theorem 4.58, so we conclude that

$$(X \otimes_R M) \otimes_R R_P \cong X \otimes_{R_P} M_P.$$

Thus

$$0 \longrightarrow A \otimes_{R_P} M_P \longrightarrow B \otimes_{R_P} M_P \longrightarrow C \otimes_{R_P} M_P \longrightarrow 0$$

is exact, and M_P is a flat R_P-module.

So whenever M is flat, M_P is a flat R_P-module for all primes P. By Theorem 4.55, M_P must be free over R_P for all primes P, that is, M is locally free.

Finally, suppose that M is locally free. We want to show that M is projective. So by Theorem 4.4, we need to show that for all surjective R-module maps $f\colon A \to B$, the map $f_*\colon \mathrm{Hom}_R(M, A) \to \mathrm{Hom}_R(M, B)$ is surjective. By Theorem 4.59, it is enough to show that f_P is surjective for all primes P. By Theorem 4.60,

$$\mathrm{Hom}_{R_P}(M_P, A_P) \cong \mathrm{Hom}_R(M, A)_P \quad \text{and} \quad \mathrm{Hom}_{R_P}(M_P, B_P) \cong \mathrm{Hom}_R(M, B)_P,$$

and

$$(f_*)_P = (f_P)_*\colon \mathrm{Hom}_{R_P}(M_P, A_P) \to \mathrm{Hom}_{R_P}(M_P, B_P).$$

But M_P i free, and thus projective by Theorem 4.3, so $(f_P)_*$ is surjective. Since this holds for all P, by Theorem 4.59 we conclude that f_* is surjective, and thus M is projective. □

Note that the noetherianity assumption is just so that finitely generated implies finitely presented; the statement is also true for a general commutative ring if instead of finitely generated modules we take finitely presented.

Exercises

Exercise 4.62. Let k be a field.

1. Show that every short exact sequence of k-vector spaces splits.
2. Prove the Rank-Nullity Theorem: given any linear transformation $T : V \to W$ of k-vector spaces,
$$\dim(\operatorname{im} T) + \dim(\ker T) = \dim V.$$

Chapter 5

Resolutions

To describe an R-module M, we need to know a set of generators and the relations among those generators. If we continue that process, and ask for relations among the relations (treating the relations as generators for the module of relations), and relations among the relations among the relations, and so on, we construct what is known as a free resolution for M. Free resolutions play a key role in many important constructions, and encode a lot of interesting information about our module. For example, if the module came from some geometric setting, geometric information about the module gets reflected in the free resolution. Studying the resolutions of *all* finitely generated modules over a ring R also tells us important information about the ring itself, and its *singularities*.

In this chapter we will introduce free resolutions, and more generally projective resolutions, as well as their injective counterpart. We will also study free resolutions in a bit more detail over commutative local noetherian rings, and the graded analogue. For more details on the basics of graded free resolutions, we recommend Irena Peeva's excellent book [Pee11].

5.1 Projective resolutions

Definition 5.1. Let M be an R-module. A **projective resolution** is a complex

$$P_\bullet = \cdots \longrightarrow P_{\underset{n}{n}} \longrightarrow \cdots \longrightarrow P_{\underset{1}{1}} \longrightarrow P_{\underset{0}{0}} \longrightarrow 0$$

where all the P_i are projective, $\mathrm{H}_0(P) = M$, and $\mathrm{H}_i(P) = 0$ for all $i \neq 0$. We may also write a projective resolution for M as an exact sequence

$$\cdots \longrightarrow P_{\underset{n}{n}} \longrightarrow \cdots \longrightarrow P_{\underset{1}{1}} \longrightarrow P_{\underset{0}{0}} \longrightarrow M \longrightarrow 0$$

where all the modules P_i are projective. The resolution is **free** if all the P_i are free.

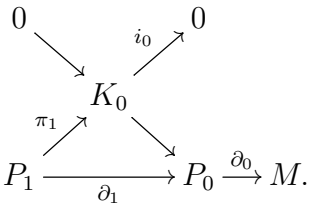
You will find both these definitions in the literature, often indicating the second option as an abuse of notation. We will be a bit sloppy and consider both equivalently, since at the end of the day they contain the same information. One often uses the word **acyclic** to refer to a complex that is exact everywhere except at homological degree 0; but we caution the reader that some authors use the word acyclic to refer to exact complexes. For that reason, we will avoid the word acyclic altogether.

Theorem 5.2. *Every R-module has a free resolution, and thus it has a projective resolution.*

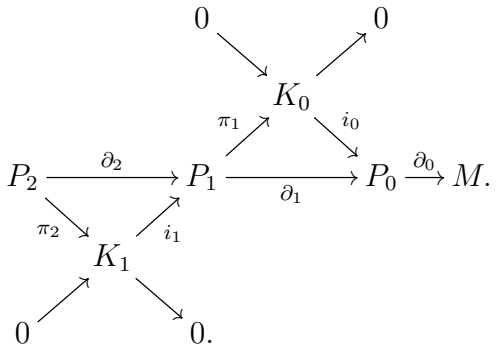
Proof. Let M be an R-module. We are going to construct a projective resolution quite explicitly. The first step is to find a projective module P_0 that surjects onto M. In fact, we can find a free module surjecting onto M, by Theorem 4.13. Now consider the kernel of that projection, say

$$0 \longrightarrow K_0 \xrightarrow{i_0} P_0 \xrightarrow{\pi_0} M \longrightarrow 0.$$

Set $\partial_0 := \pi_0$. There exists a free module P_1 surjecting onto K_0. Now the map $\partial_1 = i_0\pi_1$ satisfies $\operatorname{im}\partial_1 = K_0 = \ker\partial_0$.



Now the process continues analougously. We find a free module P_2 surjecting onto $K_1 := \ker\ \partial_1$, and set



At each stage, $\pi_i \colon P_i \longrightarrow K_{i-1}$ is a surjective map, $K_i := \ker\ \partial_i$, i_i is the inclusion of the kernel of ∂_i into P_i, and we get short exact sequences

$$0 \longrightarrow K_{n+1} \xrightarrow{i_{n+1}} P_{n+1} \xrightarrow{\pi_{n+1}} K_n \longrightarrow 0.$$

In fact, $\operatorname{im}(i_{n+1}) = \ker\partial_{n+1} = \ker(i_n\pi_{n+1}) = \ker\pi_{n+1}$. We can continue this process for as long as $P_n \neq 0$, and the resulting sequence will be a projective resolution for M. □

A free resolution

$$\cdots \longrightarrow F_2 \longrightarrow F_1 \longrightarrow F_0 \longrightarrow M$$

gives us a detailed description of our module M:

- F_0 gives us generators for M.
- F_1 gives us generators for all the relations among our generators for M.
- The next module describes the relations among the relations among our generators.

And so on.

Definition 5.3. If P is a projective resolution of M, we say that P has **length** d if $P_n = 0$ for all $n > d$ and $P_d \neq 0$. If no such d exists, we say that P has infinite length. If M has no finite projective resolution, we say that M has infinite projective dimension; otherwise, the **projective dimension** of M is the smallest length of a projective resolution.

Remark 5.4. A module M has $\operatorname{pdim}(M) = 0$ if and only M is projective. Indeed, note that if M is projective, then

$$0 \longrightarrow \underset{0}{M} \longrightarrow M \longrightarrow 0$$

is a projective resolution of M. On the other hand, if M has a projective resolution

$$0 \longrightarrow \underset{0}{P} \longrightarrow M \longrightarrow 0$$

then exactness tells us that $P \cong M$.

Example 5.5. Let us construct a free resolution for $\mathbb{Z}/2$ over $\mathbb{Z}$. First, since $\mathbb{Z}/2$ has only one generator, we can start with the canonical surjection $\pi\colon \mathbb{Z} \to \mathbb{Z}/2$. Note that $\ker \pi = (2)$ is generated by just one element again, so we can take

$$\mathbb{Z} \xrightarrow{2} \mathbb{Z} \xrightarrow{\pi} \mathbb{Z}/2.$$

But now the map $\mathbb{Z} \xrightarrow{2} \mathbb{Z}$ is injective, so we are done, and

$$0 \longrightarrow \mathbb{Z} \xrightarrow{2} \mathbb{Z} \longrightarrow \mathbb{Z}/2 \longrightarrow 0$$

is a free resolution for $\mathbb{Z}/2$. This shows that $\operatorname{pdim}(\mathbb{Z}/2) \leqslant 1$. Also, $\mathbb{Z}/2$ is not projective: we showed in Theorem 3.27 that $\operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}/2, -)$ is not exact. Thus $\operatorname{pdim}(\mathbb{Z}/2) = 1$.

Example 5.6. Consider a field k and $R = k[x]/(x^3)$. Let us construct a free resolution for $M = R/(x)$. We can start with the canonical surjection $R \to M$; the kernel is (x), which is cyclic, so our resolution begins with

$$R \xrightarrow{x} R \longrightarrow M.$$

Now the kernel of $R \xrightarrow{x} R$ is (x^2), which is again cyclic. Our resolution continues with

$$R \xrightarrow{x^2} R \xrightarrow{x} R \longrightarrow M.$$

Next, we need to compute the kernel of multiplication by x^2; but that is (x), a cyclic module, and the next step in the resolution is

$$R \xrightarrow{x} R \xrightarrow{x^2} R \xrightarrow{x} R \longrightarrow M.$$

But now we have a repeating pattern! Our two-periodic resolution goes on forever:

$$\cdots \longrightarrow R \xrightarrow{x^2} R \xrightarrow{x} R \xrightarrow{x^2} R \xrightarrow{x} R \longrightarrow M.$$

In fact, it turns out that $\operatorname{pdim}(M) = \infty$. But to really justify that, we need to understand that this is a *minimal* free resolution.

To talk about *minimal* free resolutions we need some reasonable conditions to hold. For the rest of the section, all rings will be commutative, and in fact we will be focusing on two types of rings: commutative local rings or $\mathbb{N}$-graded algebras over fields.

When k is a field, the polynomial ring $R = k[x_1, \dots, x_n]$ can be given an $\mathbb{N}$-grading by setting $\deg(x_i) = d_i$ for some $d_i \in \mathbb{N}$. The most common $\mathbb{N}$-grading, also known as the **standard grading**, is the one where we declare $\deg(x_i) = 1$ for all i. Once we declare the degrees of the variables, we can extend that grading to all monomials as follows:

$$\deg\left(x_1^{a_1} \cdots x_n^{a_n}\right) = a_1 d_1 + \cdots + a_n d_n.$$

A **homogeneous element** in R is any k-linear combination of monomials of the same degree. We write R_i for the set of all homogeneous elements of degree i, which is an abelian group under addition, and note that

$$R = \bigoplus_i R_i.$$

Note also that $R_i R_j \subseteq R_{i+j}$ for all i and j. More generally, a **graded ring** is any ring that can be decomposed in pieces of this form, meaning that

$$R = \bigoplus_i R_i \qquad \text{and} \qquad R_i R_j \subseteq R_{i+j}.$$

The elements in R_i are called homogeneous elements of degree i. Similarly, a graded R-module is a module such that

$$M = \bigoplus_i M_i \qquad \text{and} \qquad R_i M_j \subseteq M_{i+j}.$$

A homomorphism of graded R-modules $\varphi \colon M \to N$ that such that $\varphi(M_i) \subseteq N_{i+d}$ for all i is a **graded map** of degree d. Any graded map can be thought of as a map of degree 0 by shifting degrees. We write $M(-d)$ for the graded R-module with $M(-d)_i = M_{i-d}$.

When $R = k[x_1, \dots, x_n]$ is standard graded,

$$R_i = \bigoplus_{a_1 + \cdots + a_n = i} x_1^{a_1} \cdots x_n^{a_n}.$$

Note here that 0 can be though of as a homogeneous element of any degree; one sometimes declares $\deg(0) = -\infty$. An ideal I in R is a **homogeneous ideal** if it can be generated by homogeneous elements; one can show that this is equivalent to

$$I = \bigoplus_i (I \cap R_i).$$

Finally, whenever I itself is homogeneous, the grading on R passes onto R/I, with

$$(R/I)_i = R_i/I_i.$$

We will be concerned with finitely generated $\mathbb{N}$-graded k-algebras R with $R_0 = k$, which are of the form $R = k[x_1, \dots, x_n]/I$ for some homogeneous ideal I. One nice feature of such rings

is that while there might be many maximal ideals, there is only one *homogeneous* maximal ideal, which is given by

$$R_+ := \bigoplus_{i>0} R_i.$$

In many ways, the behavior of such a graded ring and its unique homogeneous maximal ideal R_+ is an analogue to the behavior of a local ring R and its unique maximal ideal ideal $\mathfrak{m}$, though one always needs to provide a separate proof for the graded and local versions.

Definition 5.7. Let $(R, \mathfrak{m})$ be either a commutative local ring or a commutative $\mathbb{N}$-graded k-algebra with $R_0 = k$ and homogeneous maximal ideal $\mathfrak{m} = R_+$. A complex

$$\cdots \longrightarrow F_2 \xrightarrow{\partial_2} F_1 \xrightarrow{\partial_1} F_0 \longrightarrow \cdots$$

is **minimal** if $\operatorname{im} \partial_{n+1} \subseteq \mathfrak{m}F_n$ for all n.

Remark 5.8. A complex (F, ∂) is minimal if and only if the differentials in the complex $F \otimes_R R/\mathfrak{m}$ are all identically 0. If all the F_i are free, fix a basis for each F_i. The differentials ∂_i can be represented by matrices, though possibly infinite. We will be primarily interested in the case of finitely generated modules over noetherian rings, which are finitely presented, so all the F_i are finitely generated as well, and each ∂_i corresponds to some finite matrix. In this case, our complex is minimal if and only if all the entries in the matrices representing ∂_i are in $\mathfrak{m}$, whatever our chosen bases are.

Lemma 5.9. *Let R be a commutative ring. Suppose $(R, \mathfrak{m})$ is either a local ring or an $\mathbb{N}$-graded k-algebra with $R_0 = k$ and homogeneous maximal ideal $\mathfrak{m} = R_+$. Let M be a finitely generated (graded) R-module. A free resolution*

$$F = \cdots \longrightarrow F_2 \xrightarrow{\partial_2} F_1 \xrightarrow{\partial_1} F_0$$

for M is a minimal complex if and only if for all n the module F_n is the free module on a minimal set of generators for $\ker \partial_{n-1}$, which in the graded case must be homogeneous.

Proof. Suppose there exists an n such that F_n is the free module on some non-minimal set of generators $m_1, \ldots, m_s$ for $K_{n-1} := \ker \partial_{n-1}$; so there is a basis $e_1, \ldots, e_s$ for F_n such that $\partial_n(e_i) = m_i$, and the images of $m_1, \ldots, m_s$ in the vector space $K_{n-1}/\mathfrak{m}K_{n-1}$ are linearly dependent. Then there exists $r_1, \ldots, r_s \in R$, not all in $\mathfrak{m}$, such that $r_1 m_1 + \cdots + r_s m_s = 0$ in R. In the graded case, we can take all these coefficients r_i to be homogeneous. At least one of these coefficients is not in $\mathfrak{m}$, and thus it must be invertible,[1] so we can multiply by its inverse. So perhaps after reordering our elements, we get

$$m_s = r_1 m_1 + \cdots + r_{s-1} m_{s-1}.$$

Then

$$e_s - r_1 e_1 - \cdots - r_{s-1} e_{s-1} \in \ker \partial_n = \operatorname{im} \partial_{n+1}$$

is not in $\mathfrak{m}F_n$, so $\operatorname{im} \partial_{n+1} \nsubseteq \mathfrak{m}F_n$.

[1] In the graded case, homogeneous elements not in $\mathfrak{m}$ are nonzero elements in $R_0 = k$, and thus invertible.

Now suppose that $\operatorname{im}\partial_{n+1} \not\subseteq \mathfrak{m}F_n$ for some n. Let $e_1, \dots, e_s$ be a basis for F_n, so that $\partial_n(e_1), \dots, \partial_n(e_s)$ form a generating set for $K_{n-1} := \ker \partial_{n-1}$. By assumption, $\ker \partial_n = \operatorname{im}\partial_{n+1}$ contains some (homogeneous, in the graded case) element that is not in $\mathfrak{m}F_n$. So there is an element $r_1e_1 + \cdots + r_se_s \in \ker \partial_n$ not in $\mathfrak{m}F_n$. In particular, some $r_i \notin \mathfrak{m}$, which we can assume without loss of generality to be r_1. Multiplying by the inverse of r_1, we get some $c_i \in R$ such that

$$e_1 - c_2e_2 - \cdots - c_se_s \in \ker \partial_n,$$

so

$$\partial_n(e_1) = c_2\partial_n(e_2) + \cdots + c_s\partial_n(e_s).$$

This is a nontrivial relation among our chosen set of generators of K_{n-1}, which must then be non-minimal. □

So to construct a minimal free resolution of M, we simply take as few generators as possible in each step. Ultimately, we can talk about *the* minimal free resolution of M. To show that, we need some definitions and a lemma.

Definition 5.10. Let (F, ∂) and (G, δ) be complexes of R-modules. The **direct sum** of F and G is the complex of R-modules $F \oplus G$ that has $(F \oplus G)_n = F_n \oplus G_n$, with differentials given by

$$\begin{array}{ccc} F_{n+1} & \xrightarrow{\partial_{n+1}} & F_n \\ \oplus & & \oplus \\ G_{n+1} & \xrightarrow[\delta_{n+1}]{} & G_n, \end{array}$$

together with the complex maps $F \to F \oplus G$ and $G \to F \oplus G$ given by the corresponding inclusion in each homological degree.

Remark 5.11. When R is a graded ring and M is a graded R-module, we can talk about *graded* direct summands of M. A module N is a **graded direct summands** of M if N is a graded R-module, there is an injective graded map $N \to M$ (of degree 0), and this inclusion splits by a graded splitting (of degree 0). This is a strictly stronger condition than simply being a direct summand.

Exercise 5.12. Show that the direct sum of complexes is the coproduct in the category $\mathrm{Ch}(R)$.

Remark 5.13. The homology of a direct sum is the direct sum of the homologies, since

$$(\partial_n, \delta_n)(a, b) = (0, 0) \iff \partial_n(a) = 0 \text{ and } \delta_n(b) = 0,$$

and

$$(a, b) \in \operatorname{im}(\partial_n, \delta_n) \text{ if and only if } a \in \operatorname{im}\partial_n \text{ and } b \in \operatorname{im}\partial_n.$$

Thus

$$\mathrm{H}_n(F \oplus G) = \frac{\ker(\partial_n, \delta_n)}{\operatorname{im}(\partial_{n+1}, \delta_{n+1})} = \frac{\ker \partial_n}{\operatorname{im}\partial_{n+1}} \oplus \frac{\ker \delta_n}{\operatorname{im}\delta_{n+1}} = \mathrm{H}_n(F) \oplus \mathrm{H}_n(G).$$

Remark 5.14. Suppose that C is a subcomplex of D, and that we know that each C_n is a direct summand of D_n, say by $D_n = C_n \oplus B_n$. In order for C to be a direct summand of D, we also need that the differentials of D behave well with C: for each n, we need to check that $\partial_n(B_n) \subseteq B_{n-1}$ and $\partial_n(C_n) \subseteq C_{n-1}$. This does not always hold.

Definition 5.15. A complex C of R-modules is **trivial** if it is a direct sum of complexes of the form

$$\cdots \longrightarrow 0 \longrightarrow R \xrightarrow{1} R \longrightarrow 0 \longrightarrow \cdots .$$

Example 5.16. The complex

$$0 \longrightarrow R \xrightarrow{\begin{pmatrix}1\\0\end{pmatrix}} R^2 \xrightarrow{\begin{pmatrix}0 & 1\end{pmatrix}} R \longrightarrow 0 \qquad = \qquad \begin{array}{ccccccccc} 0 & \longrightarrow & R & \xrightarrow{1} & R & \longrightarrow & 0 \\ & & & \oplus & & & \\ & & 0 & \longrightarrow & R & \xrightarrow{1} & R & \longrightarrow & 0 \end{array}$$

is trivial.

Remark 5.17. Trivial complexes are exact: they are the direct sums of exact complexes, and by Theorem 5.13 taking homology commutes with direct sums.

Lemma 5.18. *Let $(R, \mathfrak{m})$ be either a commutative local ring or a commutative $\mathbb{N}$-graded k-algebra with $R_0 = k$ and homogeneous maximal ideal $\mathfrak{m} = R_+$. Every (graded) complex*

$$\cdots \longrightarrow T_2 \xrightarrow{\partial_2} T_1 \xrightarrow{\partial_1} T_0 \longrightarrow 0$$

of finitely generated (graded) free R-modules that is exact everywhere must be trivial.

Proof. Since T_0 is projective, Theorem 4.6 says that the short exact sequence

$$0 \longrightarrow \ker \partial_1 \longrightarrow T_1 \xrightarrow{\partial_1} T_0 \longrightarrow 0$$

splits, so $T_1 \cong \ker \partial_1 \oplus T_0$. In fact, ∂_1 is the canonical projection map $T_0 \oplus \ker \partial_1 \to T_0$, and our original exact sequence breaks off as

$$\begin{array}{ccccccccc} \cdots & \longrightarrow & T_2 & \xrightarrow{\partial_2} & \ker \partial_1 & \longrightarrow & 0 \\ & & & & \oplus & & \\ & & 0 & \longrightarrow & T_0 & \xrightarrow{1} & T_0 & \longrightarrow & 0. \end{array}$$

In particular, since $0 \longrightarrow T_0 \xrightarrow{1} T_0 \longrightarrow 0$ is trivial and homology commutes with taking direct sums of complexes, by Theorem 5.13, we conclude that

$$\cdots \longrightarrow T_2 \xrightarrow{\partial_2} \ker \partial_1 \longrightarrow 0$$

is also exact everywhere. In particular, we have also shown that $\ker \partial_1$ is a (graded) direct summand of the (graded) free R-module T_1. In the local case, $\ker \partial_1$ is projective by Theorem 4.9, and thus free by Theorem 4.55. In the graded setting, one can also show that

$\ker \partial_1$ is free. So we are back at our original situation, and we can repeat the same argument repeatedly to show that our complex breaks off as the direct sum of the trivial complexes

$$0 \longrightarrow \ker \partial_n \xrightarrow{1} \ker \partial_n \longrightarrow 0$$

and must therefore be trivial. □

Theorem 5.19. *Consider a complex*

$$P = \cdots \longrightarrow P_n \longrightarrow \cdots P_1 \xrightarrow{\partial_1} P_0 \xrightarrow{\partial_0} M \longrightarrow 0$$

with all the P_i projective R-modules, and let

$$C = \cdots \longrightarrow C_n \longrightarrow \cdots C_1 \xrightarrow{\delta_1} C_0 \xrightarrow{\delta_0} N \longrightarrow 0$$

be any exact complex. Every R-module map $f\colon M \to N$ can be lifted to a map of complexes $\varphi\colon P \to C$, and any two such lifts are homotopic.

Moreover, if R is a commutative graded k-algebra, M and N are finitely generated graded R-modules, P_n and C_n are finitely generated graded R-modules, and f is a degree-preserving homomorphism, then the induced map of complexes is made out of degree-preserving R-module maps.

Proof. Since P_0 is projective and δ_0 is surjective, there exists an R-module homomorphism φ_0 such that

$$\begin{array}{ccccc} P_0 & \xrightarrow{\partial_0} & M & \longrightarrow & 0 \\ {\scriptstyle \varphi_0}\downarrow & & \downarrow{\scriptstyle f} & & \\ C_0 & \xrightarrow[\delta_0]{} & N & \longrightarrow & 0 \end{array}$$

commutes. Notice in fact that

$$\begin{aligned} \delta_0\varphi_0(\operatorname{im}\partial_1) &\subseteq \delta_0\varphi_0(\ker\partial_0) && \text{because } P \text{ is a complex} \\ &= f\partial_0(\ker\partial_0) && \text{by commutativity of the square above} \\ &= 0, \end{aligned}$$

so $\varphi_0(\operatorname{im}\partial_1) \subseteq \ker\delta_0 = \operatorname{im}\delta_1$. In the graded case, note that we can define φ_0 by sending the elements b_i in a homogeneous basis of P_0 to homogeneous $c_i \in C_0$ such that $\delta_0(c_i) = f\partial_0(b_i)$.

We now proceed by induction. Suppose we have constructed $P_{n-1} \xrightarrow{\varphi_{n-1}} C_{n-1}$ such that $\varphi_{n-1}(\operatorname{im}\partial_n) \subseteq \operatorname{im}\delta_n$. Since P_n is projective, there exists a map φ_n such that

$$\begin{array}{ccc} P_n & \xrightarrow{\partial_n} & P_{n-1} \\ {\scriptstyle \varphi_n}\downarrow & & \downarrow{\scriptstyle \varphi_{n-1}} \\ C_n & \xrightarrow[\delta_n]{} & \operatorname{im}\delta_n \end{array}$$

commutes. And again,

$$\begin{aligned}
\delta_n\varphi_n(\operatorname{im}\partial_{n+1}) &\subseteq \delta_n\varphi_n(\ker\partial_n) && \text{because } P \text{ is a complex}\\
&= \varphi_{n-1}\partial_n(\ker\partial_n) && \text{by commutativity of the square above}\\
&= 0,
\end{aligned}$$

so $\varphi_n(\operatorname{im}\partial_{n+1}) \subseteq \ker\delta_n = \operatorname{im}\delta_{n+1}$.

We can now inductively construct our map of complexes φ lifting f.

Now suppose we are given two such maps of complexes $P \longrightarrow C$ lifting f, say φ and ψ. Note that $\varphi - \psi$ and 0 are two liftings of the zero map. We are going to show that any map lifting the zero map $M \longrightarrow N$ must be nullhomotopic, which will then imply that φ and ψ are homotopic as well (essentially via the same homotopy!).

So let $\varphi\colon P \longrightarrow C$ be a map of complexes lifting the zero map $M \longrightarrow N$, so that the following commutes:

$$\begin{array}{ccccccccc}
\cdots & \longrightarrow & P_1 & \xrightarrow{\partial_1} & P_0 & \xrightarrow{\partial_0} & M & \longrightarrow & 0\\
 & & \downarrow{\scriptstyle\varphi_1} & & \downarrow{\scriptstyle\varphi_0} & & \downarrow{\scriptstyle 0} & & \\
\cdots & \longrightarrow & C_1 & \xrightarrow{\delta_1} & C_0 & \xrightarrow{\delta_0} & N & \longrightarrow & 0
\end{array}$$

We will explicitly construct a nullhomotopy for φ by induction. First, set $h_n = 0$ for all $n < 0$. The commutativity of the rightmost square tells us that $\delta_0\varphi_0 = 0$, so

$$\operatorname{im}\varphi_0 \subseteq \ker\delta_0 = \operatorname{im}\delta_1.$$

Since P_0 is projective, there exists an R-module homomorphism h_0 such that

$$\begin{array}{ccc}
 & & P_0\\
 & \overset{h_0}{\swarrow} & \downarrow{\scriptstyle\varphi_0}\\
C_1 & \xrightarrow[\delta_1]{} & \operatorname{im}\delta_1
\end{array}$$

commutes, and thus $\varphi_0 = \delta_1 h_0 = \delta_1 h_0 + h_{-1}\partial_0$. Notice also that

$$\begin{aligned}
\delta_1(\varphi_1 - h_0\partial_1) &= \varphi_0\partial_1 - \delta_1 h_0\partial_1 && \text{because } \varphi \text{ is a map of complexes}\\
&= (\varphi_0 - \delta_1 h_0)\partial_1 && \text{factoring}\\
&= 0 && \text{since } \varphi_0 = \delta_1 h_0,
\end{aligned}$$

so $\operatorname{im}(\varphi_1 - h_0\partial_1) \subseteq \ker\delta_1 = \operatorname{im}\delta_2$.

Now assume that we have constructed maps $h_0, \ldots, h_n$ such that $\varphi_n = h_{n-1}\partial_n + \delta_{n+1}h_n$ and $\operatorname{im}(\varphi_{n+1} - h_n\partial_{n+1}) \subseteq \operatorname{im}\delta_{n+2}$. Since P_{n+1} is projective, we can find a map h_{n+1} such that

$$\begin{array}{ccc}
 & & P_{n+1}\\
 & \overset{h_{n+1}}{\swarrow} & \downarrow{\scriptstyle\varphi_{n+1}-h_n\partial_{n+1}}\\
C_{n+2} & \xrightarrow[\delta_{n+2}]{} & \operatorname{im}\delta_{n+2}
\end{array}$$

commutes, so $\varphi_{n+1} = \delta_{n+2}h_{n+1} + h_n\partial_{n+1}$. Now

$$\begin{aligned}
\delta_{n+2}(\varphi_{n+2} - h_{n+1}\partial_{n+2}) &= \varphi_{n+1}\partial_{n+2} - \delta_{n+2}h_{n+1}\partial_{n+2} && \text{since } \varphi \text{ is a map of complexes} \\
&= (\varphi_{n+1} - \delta_{n+2}h_{n+1})\partial_{n+2} \\
&= h_n\partial_{n+1}\partial_{n+2} && \text{by commutativity of the triangle above} \\
&= 0 && \text{since } \partial_{n+1}\partial_{n+2} = 0.
\end{aligned}$$

So we again obtain $\operatorname{im}(\varphi_{n+2} - h_{n+1}\partial_{n+2}) \subseteq \ker \delta_{n+1} = \operatorname{im} \delta_{n+2}$. By induction, this process allows us to construct our homotopy h. □

Theorem 5.20. *Let $(R, \mathfrak{m})$ be a commutative noetherian ring, which is either a local ring or an $\mathbb{N}$-graded graded k-algebra with $R_0 = k$ and homogeneous maximal ideal $\mathfrak{m} = R_+$. If F is a minimal free resolution of M, then any free resolution for M is isomorphic to a direct sum of F with a trivial complex. In particular, the minimal free resolution of M is unique up to isomorphism.*

Proof. Suppose that G is another free resolution of M. By Theorem 5.19, there are complex maps $\psi\colon G \longrightarrow F$ and $\varphi\colon F \longrightarrow G$ that lift the identity map on M. Then $\psi\varphi\colon F \longrightarrow F$ is a map of complexes that lifts the identity on M, and thus by Theorem 5.19 $\varphi\psi$ must be homotopic to the identity on F. Let h be a homotopy between $\psi\varphi$ and the identity, so that for all n,

$$\mathrm{id} - \psi_n\varphi_n = \partial_{n+1}h_n + h_{n-1}\partial_n.$$

Since F is minimal, we have $\operatorname{im}\partial_n \subseteq \mathfrak{m}F_{n-1}$ and $\operatorname{im}\partial_{n+1} \subseteq \mathfrak{m}F_n$, so $\operatorname{im}(\mathrm{id} - \psi_n\varphi_n) \subseteq \mathfrak{m}F_n$ for all n. Our first goal will be to show that $\psi\varphi$ is an isomorphism.

First we do the local case. Let A be the matrix representing $\psi_n\varphi_n$ in some fixed basis for F_n, and note that $\mathrm{id} - \psi_n\varphi_n$ is represented by $\mathrm{Id} - A$, so all the entries in $\mathrm{Id} - A$ must be in $\mathfrak{m}$. Our matrix A can be written as

$$A = \begin{pmatrix} 1 + a_{11} & a_{12} & & \cdots & a_{1s} \\ a_{21} & 1 + a_{22} & & \cdots & a_{2s} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{s1} & \cdots & & a_{s\,s-1} & 1 + a_{ss} \end{pmatrix}$$

for some $a_{ij} \in \mathfrak{m}$, so that $\det(A) = 1 + a$ for some $a \in \mathfrak{m}$. In particular, $\det(A)$ is invertible, and $\psi_n\varphi_n$ is an isomorphism.

In the graded case, we have to be a bit more careful: not all elements that are not in $\mathfrak{m}$ are invertible, this is only true for *homogeneous* elements. First, we fix a basis of homogeneous elements $f_1, \ldots, f_s$ for F_n with $\deg(f_1) \leqslant \deg(f_2) \leqslant \cdots \leqslant \deg(f_s)$, and set $\Phi := \mathrm{id} - \psi_n\varphi_n$. Since our map Φ is degree-preserving, $\Phi(f_i)$ is homogeneous for each i, and so we can write $\Phi(f_i)$ as a linear combination of our basis elements $f_1, \ldots, f_s$ using only pieces of degree $\deg(\Phi(f_i))$. We obtain a matrix $C = (c_{ij})$ such that $c_{ij} \neq 0 \implies \deg(c_{ij}) = \deg(f_j) - \deg(f_i)$, and C represents Φ, meaning $\Phi(f_i) = c_{i1}f_1 + \cdots + c_{is}f_s$ for all i. Now all the entries of $C = \mathrm{Id} - A$ must be in $\mathfrak{m}$, so in particular we must have $a_{ii} = 1$ for all i. Moreover, since we chose our basis to have increasing degrees, $\deg(c_{ij}) = 0$ whenever $i < j$. Since we must also have $c_{ij} \in \mathfrak{m}$ whenever $i \neq j$, we conclude that $c_{ij} = 0$ for $i < j$. We conclude that A is an upper triangular matrix. Finally, $\det(A) = a_{11} \cdots a_{ss} = 1$, and A is invertible.

So we have shown in both cases that $\psi_n\varphi_n$ is an isomorphism for all n. By Theorem 2.4, $\psi\varphi$ is in fact an isomorphism of complexes, so let $\xi\colon F \to F$ be its inverse. Now we want to claim that φ splits as a map of complexes. Notice that

$$(\xi\psi)\varphi = \xi(\psi\varphi) = \mathrm{id}_F,$$

so let us take $\xi\psi$ to be our proposed splitting for φ. Note that $(\xi\psi)_n\varphi_n = \mathrm{id}_n$ implies that our map $\xi\psi$ provides splittings for the R-module maps in each degree, by Theorem 2.26, so $G_n = \varphi_n(F_n) \oplus \ker(\xi_n\psi_n)$. we just need to prove that this splitting holds as complexes, that is, that $G = \varphi(F) \oplus \ker(\xi\psi)$ as complexes. So let $K := \ker(\xi\psi)$, and denote the differential in G by δ. We need to check that $\delta(\varphi(F)) \subseteq \varphi(F)$ and $\delta(K) \subseteq K$.

Since φ is a map of complexes, $\delta\varphi = \varphi\partial$, so we do get $\delta(\varphi(F)) \subseteq \varphi(F)$. Given $a \in K_{n+1}$, we can write $\delta_{n+1}(a) = \varphi(b) + c$ for some $b \in F_n$ and K_n, since $G_n = \varphi(F_n) \oplus K_n$. Then

$$\begin{aligned}
b &= \mathrm{id}(b) && \\
&= \xi_n\psi_n\varphi_n(b) && \text{since } \xi_n\psi_n \text{ is a splitting for } \varphi_n \\
&= \xi_n\psi_n(\varphi_n(b) + c) && \text{since } c \in K_n \\
&= \xi_n\psi_n\delta_{n+1}(a) && \text{by assumption} \\
&= \xi_n\delta_{n+1}\psi_n(a) && \text{since } \psi \text{ is a map of complexes} \\
&= \delta_{n+1}(\xi_n\psi_n)(a) && \text{since } \xi \text{ is a map of complexes} \\
&= 0 && \text{since } a \in K_n.
\end{aligned}$$

We conclude that $\delta_{n+1}(a) \in K_n$, and $\delta(K) \subseteq K$. We have now shown that $G \cong F \oplus K$.

Finally, we are going to show that K is a trivial complex. First, we claim that K_n is free for all n. We have already shown that K_n is a (graded) direct summand of a (graded) free module. In the local case, Theorem 4.9 says that K_n is projective, and then Theorem 4.55 says that K_n must in fact be free. In the graded setting, one can show that any graded module which is a direct sum of a finitely generated graded R-module is a graded free module. In both cases, K_n is free.

Since $G \cong F \oplus K$, we have $\mathrm{H}_n(G) \cong \mathrm{H}_n(F) \oplus \mathrm{H}_n(K)$. Since F and G are both (graded) free resolutions for M, they have the same homology: $\mathrm{H}_n(F) = \mathrm{H}_n(G) = 0$ for all $n \neq 0$, and $\mathrm{H}_0(F) = \mathrm{H}_0(G) = M$. We conclude that K is exact everywhere. Finally, Theorem 5.18 shows that K is trivial. □

Theorem 5.21 (Horseshoe Lemma)**.** *Consider a short exact sequence of modules*

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0.$$

Let P be a projective resolution of A, and R be a projective resolution of C. There exists a projective resolution Q of B and maps of complexes F and G lifting f and g such that

$$0 \longrightarrow P \xrightarrow{F} Q \xrightarrow{G} R \longrightarrow 0$$

is a short exact sequence of complexes.

Proof. First, we need to introduce some general notation: given homomorphisms $f\colon M \to L$ and $g\colon N \to L$ with the same target, we will write $f \oplus g$ for the homomorphism $M \oplus N \to L$ given by $(f \oplus g)(m, n) = f(m) + g(n)$. Moreover, we will denote the differential of P by ∂^P, and the differential of R by ∂^R.

For each $n \geqslant 0$, set $Q_n := P_n \oplus R_n$, and let $F_n : P_n \longrightarrow Q_n$ and $G_n : Q_n \longrightarrow R_n$ be the canonical projections. By Theorem 4.11, Q_n is projective for all n. Moreover, by construction we get short exact sequences

$$0 \longrightarrow P_n \xrightarrow{F_n} Q_n \xrightarrow{G_n} R_n \longrightarrow 0$$

for all n. We will construct the missing differentials ∂^Q inductively.

Since R_0 is projective and g is surjective, there exists γ such that

$$\begin{array}{ccccccccc} 0 & \longrightarrow & P_0 & \xrightarrow{F_0} & Q_0 & \xrightarrow{G_0} & R_0 & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \partial_0} & & & \swarrow{\scriptstyle \gamma} & \downarrow{\scriptstyle \partial_0} & & \\ 0 & \longrightarrow & A & \xrightarrow[f]{} & B & \xrightarrow[g]{} & C & \longrightarrow & 0 \end{array}$$

commutes. Set $\partial_0^Q := (f\partial_0^P) \oplus \gamma$. The universal property of the coproduct guarantees that

$$\begin{array}{ccccccccc} 0 & \longrightarrow & P_0 & \xrightarrow{F_0} & Q_0 & \xrightarrow{G_0} & R_0 & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \partial_0} & & \downarrow{\scriptstyle \partial_0^Q} & \swarrow{\scriptstyle \gamma} & \downarrow{\scriptstyle \partial_0} & & \\ 0 & \longrightarrow & A & \xrightarrow[f]{} & B & \xrightarrow[g]{} & C & \longrightarrow & 0 \end{array}$$

commutes. By the Five Lemma, ∂_0^Q is surjective. Moreover, ∂_0^P is also surjective, so $\operatorname{coker}(\partial_0^P) = 0$. By the Snake Lemma,

$$0 \longrightarrow \ker \partial_0^P \longrightarrow \ker \partial_0^Q \longrightarrow \ker \partial_0^R \longrightarrow \operatorname{coker}(\partial_0^P) = 0$$

is exact. We then proceed by induction, and at each step we apply the base case to the commutative diagram with exact rows

$$\begin{array}{ccccccccc} 0 & \longrightarrow & P_{n+1} & \xrightarrow{F_{n+1}} & Q_{n+1} & \xrightarrow{G_{n+1}} & R_n & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \partial_{n+1}^P} & & & & \downarrow{\scriptstyle \partial_{n+1}^R} & & \\ 0 & \longrightarrow & \ker \partial_n^P & \longrightarrow & \ker \partial_n^Q & \longrightarrow & \ker \partial_n^R & \longrightarrow & 0 \end{array}$$

where the vertical arrows are surjective because P and R are projective resolutions and thus exact. Notice that by construction, the image of ∂_{n+1}^Q is contained in $\ker \partial_n^Q$, which guarantees that ∂ is a differential.

This inductive process allows us to build a complex of projectives Q and a short exact sequence of complexes

$$0 \longrightarrow P \xrightarrow{F} Q \xrightarrow{G} R \longrightarrow 0.$$

Applying the long exact sequence in homology, we get exact sequences

$$0 = \mathrm{H}_n(P) \longrightarrow \mathrm{H}_n(Q) \longrightarrow \mathrm{H}_n(R) = 0$$

for all $n \geqslant 1$, and thus $\mathrm{H}_n(Q) = 0$. Moreover, we constructed δ_0^Q so that

$$Q_1 \xrightarrow{\delta_1} Q_0 \longrightarrow B$$

is exact, and thus $\mathrm{H}_0(Q) = B$. We conclude that Q is a projective resolution of B. □

Now that we know that minimal free resolutions exist and are unique (in the local and graded settings), we will take the rest of this section to briefly discuss how minimal free resolutions contain a lot of important information about our modules. For example, we want to keep track of the kernels of the differentials in a minimal free resolution.

Definition 5.22. Let $(R, \mathfrak{m})$ be a commutative ring, either a local ring or an $\mathbb{N}$-graded k-algebra with $R_0 = k$ and homogeneous maximal ideal $\mathfrak{m} = R_+$. Let F be a minimal free resolution for the finitely generated (graded) R-module M. For each $n \geqslant 1$, the submodule

$$\Omega_n(M) := \operatorname{im} \partial_n = \ker \partial_{n-1}$$

is the n**th syzygy** of M.

Remark 5.23. For each n, we have a short exact sequence

$$0 \longrightarrow \ker \partial_n \longrightarrow F_n \longrightarrow \operatorname{im} \partial_n \longrightarrow 0.$$

But $\ker \partial_n = \Omega_{n+1}(M)$ and $\operatorname{im} \partial_n = \Omega_n(M)$, so we get a short exact sequence

$$0 \longrightarrow \Omega_{n+1}(M) \longrightarrow F_n \longrightarrow \Omega_n(M) \longrightarrow 0.$$

Syzygies are indeed well-defined up to isomorphism.

Remark 5.24. Suppose that F and G are two minimal free resolutions for M. By Theorem 5.20, there exists an isomorphism between F and G, say φ. Since φ is a map of complexes, $\varphi\partial^F = \partial^G\varphi$, and thus φ must send elements in $\ker \partial^F$ into elements in $\ker \partial^G$. Similarly, an inverse ψ to φ sends $\ker \partial^G$ into $\ker \partial^F$. In each homological degree, the induced maps $\ker \partial_n^F \longrightarrow \ker \partial_n^G$ and $\ker \partial_n^F \longrightarrow \ker \partial_n^G$ are inverse, and thus isomorphisms. In the graded case, one can show that we obtain graded isomorphisms, so that the graded syzygies are also well-defined up to isomorphism.

The number of generators in each homological degree is also an important invariant.

Definition 5.25. Let $(R, \mathfrak{m})$ be a commutative ring, either a local ring or an $\mathbb{N}$-graded k-algebra with $R_0 = k$ and homogeneous maximal ideal $\mathfrak{m} = R_+$. Let F be a minimal free resolution for the finitely generated (graded) R-module M. The nth **betti number** of M is

$$\beta_i(M) := \operatorname{rank} F_i = \mu(F_i).$$

In the graded case, we can also talk about *graded* betti numbers. When M is a graded module, we can write a resolution that keeps track of the grading.

Definition 5.26. Let R be a commutative $\mathbb{N}$-graded graded k-algebra with $R_0 = k$ and homogeneous maximal ideal $\mathfrak{m} = R_+$. Let M be a graded R-module. The (i,j)th betti number of M, $\beta_{ij}(M)$, counts the number of generators of F_i in degree j. We often collect the betti numbers of a module in its **betti table**:

$\beta(M)$	0	1	2	$\cdots$
0	$\beta_{00}(M)$	$\beta_{01}(M)$	$\beta_{02}(M)$	
1	$\beta_{11}(M)$	$\beta_{12}(M)$	$\beta_{13}(M)$	
2	$\beta_{22}(M)$	$\beta_{23}(M)$		
$\vdots$			$\ddots$	

By convention, the entry corresponding to (i,j) in the betti table of M contains $\beta_{i,i+j}(M)$, and *not* $\beta_{ij}(M)$. This is how Macaulay2 displays betti tables.

Example 5.27. Let $R = k[x,y,z]$ and $M = R/(xy, xz, yz)$. The minimal free resolution for M is

$$0 \longrightarrow R^2 \xrightarrow{\begin{pmatrix} z & 0 \\ -y & y \\ 0 & -x \end{pmatrix}} R^3 \xrightarrow{\begin{pmatrix} xy & xz & yz \end{pmatrix}} R \longrightarrow M.$$

From this minimal resolution, we can read the betti numbers of M:

- $\beta_0(M) = 1$, since M is a cyclic module;
- $\beta_1(M) = 3$, and these three quadratic generators live in degree 2;
- $\beta_2(M) = 2$, and these represent linear syzygies on quadrics, and thus live in degree 3.

To write a graded free resolution for M, we choose all maps to have degree 0, so that the graded free modules in each degree are sums of copies of shifts of R. Here is the graded free resolution of M:

$$0 \longrightarrow R(-3)^2 \xrightarrow{\begin{pmatrix} z & 0 \\ -y & y \\ 0 & -x \end{pmatrix}} R(-2)^3 \xrightarrow{\begin{pmatrix} xy & xz & yz \end{pmatrix}} R \longrightarrow M.$$

Notice that the graded shifts in lower homological degrees affect all the higher homological degrees as well. For example, when we write the map in degree 2, we only need to shift the degree of each generator by 1, but since our map now lands on $R(-2)^3$, we have to bump up degrees from 2 to 3, and write $R(-3)^2$. The graded betti number $\beta_{ij}(M)$ of M counts the number of copies of $R(-j)$ in homological degree i in our resolution. So we have

$$\beta_{00} = 1, \beta_{12} = 3, \text{and } \beta_{23} = 2.$$

We can collect the graded betti numbers of M in its betti table:

$\beta(M)$	0	1	2
0	1	–	–
1	–	3	2

.

Example 5.28. Let k be a field, $R = k[x, y]$, and consider the ideal

$$I = (x^2, xy, y^3)$$

which has two generators of degree 2 and one of degree 3, so there are graded betti numbers β_{12} and β_{13}. The minimal free resolution for R/I is

$$0 \longrightarrow \begin{matrix} R(-3)^1 \\ \bigoplus \\ R(-4)^1 \end{matrix} \xrightarrow{\begin{pmatrix} y & 0 \\ -x & y^2 \\ 0 & -x \end{pmatrix}} \begin{matrix} R(-2)^2 \\ \bigoplus \\ R(-3)^1 \end{matrix} \xrightarrow{\begin{pmatrix} x^2 & xy & y^3 \end{pmatrix}} R \longrightarrow R/I.$$

$$\beta_{23}(R/I) = 1 \qquad \beta_{12}(R/I) = 2$$
$$\beta_{24}(R/I) = 1 \qquad \beta_{13}(R/I) = 1$$

So the betti table of R/I is

$\beta(M)$	0	1	2
0	1	–	–
1	–	2	1
2	–	1	1

.

In fact, even if all we know is the betti numbers of M, there is lots of information to we can extract about M. For more about the beautiful theory of free resolutions and syzygies, see [Eis05]. For a detailed treatment of graded free resolutions, see [Pee11].

5.2 Injective resolutions

Injective resolutions are analogous to projective resolutions, but now we want to approximate our module M by injectives.

Definition 5.29. Let M be an R-module. An **injective resolution** of M is a complex

$$E = \; 0 \longrightarrow E_0 \longrightarrow E_1 \longrightarrow E_2 \longrightarrow \cdots$$

with each E_i injective, $\mathrm{H}_0(E) = M$, and $\mathrm{H}_n(E) = 0$ for all $n \neq 0$. We may abuse notation and instead say that an injective resolution of M is an exact sequence

$$0 \longrightarrow M \longrightarrow E_0 \longrightarrow E_1 \longrightarrow E_2 \longrightarrow \cdots.$$

Remark 5.30. This is the first example we have encountered where we have a *cocomplex* rather than a complex. Its homology should technically be referred to as cohomology, and written with superscripts:

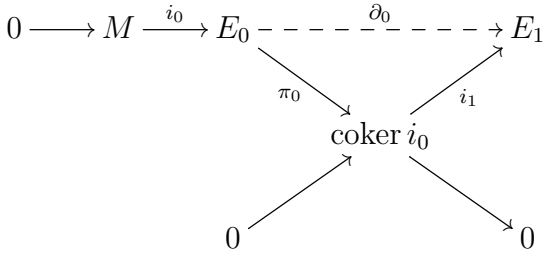
We can construct injective resolutions in a similar fashion to how we constructed projective resolutions.

Theorem 5.31. *Every R-module M has an injective resolution.*

Proof. By Theorem 4.33, every R-module embeds into an injective module. So we start by taking an injective R-module E_0 containing M, and look at the cokernel of the inclusion.

$$0 \longrightarrow M \xrightarrow{i_0} E_0 \xrightarrow{\pi_0} \operatorname{coker} i_0 \longrightarrow 0.$$

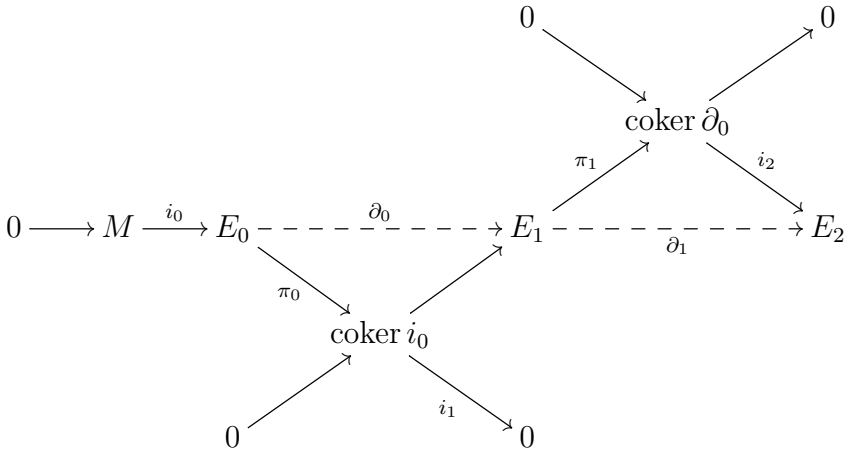
Now $\operatorname{coker} i_0$ includes in some other injective module E_1.



Take $\partial_0 := i_1\pi_0$. Since i_1 is injective,

$$\ker \partial_0 = \ker(i_1\pi_0) = \ker \pi_0 = \operatorname{im} i_0.$$

Notice also that $\operatorname{coker} i_0 = \operatorname{im} \partial_0 = \ker(E_1 \longrightarrow \operatorname{coker} \partial_0)$. So we can now we continue in a similar fashion, by finding an injective module E_2 that $\operatorname{coker} \partial_0$ embeds into.



By construction and since i_2 is injective, $\ker \partial_1 = \operatorname{im} \partial_0$, and our complex is exact at E_1. The process continues analogously. □

We can again define a minimal injective resolution for M as one where at each step we take the smallest injective module that $\operatorname{coker} i_n$ embeds into; this is called the injective hull of M. Perhaps unsurprisingly, one can show that the minimal injective resolution of a finitely generated module over a local ring is unique up to isomorphism. The analogues to the betti numbers are called Bass numbers, although now there are some major differences. When we construct a minimal free resolution, we have only to count copies of R in each homological degree, while there are many different building blocks for injective modules — the injective hulls of R/P, where P ranges over the prime ideals in R. So for each homological degree i, we get one bass number for each prime ideal P.

Example 5.32. Let's construct a minimal free resolution for the abelian group $\mathbb{Z}$. We start by including $\mathbb{Z}$ in $\mathbb{Q}$, and then note that the cokernel $\mathbb{Q}/\mathbb{Z}$ is actually injective, by Theorem 4.28 and Theorem 4.24. So $\mathbb{Q}/\mathbb{Z}$ embeds in itself, and our resolution stops there. So the short exact sequence

$$0 \longrightarrow \mathbb{Z} \longrightarrow \mathbb{Q} \longrightarrow \mathbb{Q}/\mathbb{Z} \longrightarrow 0$$

is in fact a minimal injective resolution for $\mathbb{Z}$.

Chapter 6

Derived functors

While Hom and tensor are not exact functors, we can measure their lack of exactness using their derived functors Ext and Tor. These are the poster child examples of what are called derived functors, which can be constructed over any abelian category provided we have enough projective or injective objects. In this chapter, we will construct derived functors over R-**Mod** (which does have enough injectives and enough projectives), and then later we will discuss the general construction.

6.1 The general construction

We start with the general construction of derived functors, although we will soon focus on concrete examples, most importantly Ext and Tor, the derived functors of hom and tensor.

In the following definition, when we say resolution we mean the non-augmented resolution; so a projective resolution for M is a complex P of projectives with $\mathrm{H}_0(P) = M$, $\mathrm{H}_i(P) = 0$ for all $i \neq 0$, and $P_i = 0$ for all $i < 0$, while an injective resolution for M is a cocomplex E of injectives with $\mathrm{H}^0(E) = M$, $\mathrm{H}^i(P) = 0$ for all $i \neq 0$, and $P_i = 0$ for all $i < 0$.

Definition 6.1 (Derived functors)**.** Let $F\colon R\text{-}\mathbf{Mod} \longrightarrow S\text{-}\mathbf{Mod}$ be a covariant right exact functor. The **left derived functors** of F are a sequence of functors

$$L_iF\colon R\text{-}\mathbf{Mod} \longrightarrow S\text{-}\mathbf{Mod}, \quad \text{for } i \geqslant 0,$$

defined as follows:

- For each R-module A, fix a projective resolution P of A, and set
$$L_iF(A) := \mathrm{H}_i(F(P)).$$
- Given a R-module homomorphism $f\colon A \to B$, fix projective resolutions P of A and Q of B, and a map of complexes $\varphi\colon P \to Q$ lifting f. Then
$$L_iF(f) := \mathrm{H}_i(F(\varphi)).$$

Let $F : R\text{-}\mathbf{Mod} \longrightarrow S\text{-}\mathbf{Mod}$ be a covariant left exact functor. The **right derived functors** of F are a sequence of functors

$$R^iF\colon R\text{-}\mathbf{Mod} \longrightarrow S\text{-}\mathbf{Mod}, \quad \text{for } i \geqslant 0,$$

defined as follows:

- For each R-module A, fix an injective resolution E of A, and set

$$R^iF(A) := \mathrm{H}^i(F(E)).$$

- Given an R-module homomorphism $f\colon A \to B$, fix injective resolutions E of A and I of B, and a map of complexes $\varphi\colon P \to Q$ extending f. Then

$$R^iF(f) := \mathrm{H}^i(F(\varphi)).$$

Let $F\colon R\text{-}\mathbf{Mod} \longrightarrow S\text{-}\mathbf{Mod}$ be a contravariant left exact functor. The **right derived functors** of F are a sequence of functors

$$R^iF\colon R\text{-}\mathbf{Mod} \longrightarrow S\text{-}\mathbf{Mod}, \quad \text{for } i \geqslant 0,$$

defined as follows:

- For each R-module A, fix a projective resolution P of A, and set

$$R^iF(A) := \mathrm{H}^i(F(P)).$$

- Given an R-module homomorphism $f\colon A \to B$, fix projective resolutions P for A and Q for B, and a map of complexes $\varphi\colon P \to Q$ extending f. Then

$$R^iF(f) := \mathrm{H}^i(F(\varphi)).$$

Finally, let $F : R\text{-}\mathbf{Mod} \longrightarrow S\text{-}\mathbf{Mod}$ be a contravariant right exact functor. The **left derived functors** of F are a sequence of functors

$$L_iF\colon R\text{-}\mathbf{Mod} \longrightarrow S\text{-}\mathbf{Mod}, \quad \text{for } i \geqslant 0,$$

defined as follows:

- For each object A in $\mathcal{A}$, fix an injective resolution E of A, and set

$$L_iF(A) := \mathrm{H}_i(F(E)).$$

- Given an arrow $A \xrightarrow{f} B$, fix injective resolutions $A \longrightarrow E$ and $B \longrightarrow I$, and a map of complexes $E \xrightarrow{\varphi} I$ extending f. Then

$$L_iF(f) := \mathrm{H}_i(F(\varphi)).$$

It is not clear a priori that this construction is well-defined, but we will soon show that is indeed the case.

Remark 6.2. If F is exact, then $\mathrm{H}_i(F(C)) = F(\mathrm{H}_i(C))$, by Theorem 3.17, so $L_iF = 0$ for all $i > 0$.

Remark 6.3. If P is projective, then $0 \to P \to 0$ is a projective resolution of P, and thus $L_iF(P) = 0$ for all $i > 0$. Similarly, if E is injective then $R^iF(E) = 0$.

Proposition 6.4. *Let $F\colon R$-**Mod** $\longrightarrow S$-**Mod** be a covariant right exact functor.*

a) $L_iF(A)$ is well-defined up to isomorphism for every object A.

b) $L_iF(f)$ is well-defined for every arrow f.

c) L_iF is an additive functor for each i.

d) $L_0F = F$.

Proof.

a) Let P and Q be projective resolutions of A. Theorem 5.19 gives us maps of complexes $\varphi\colon P \to Q$ and $\psi\colon Q \to P$ such that $\varphi\psi$ is homotopic to 1_Q and $\psi\varphi$ is homotopic to 1_P. Additive functors preserve homotopies, by Theorem 7.47, so $F(\varphi)F(\psi)$ and $F(\psi)F(\varphi)$ are homotopic to the corresponding identity maps. Homotopic maps induce the same map in homology, by Theorem 2.12. Therefore, $F(\varphi)$ and $F(\psi)$ induce isomorphisms in homology.

b) Fix projective resolutions P and Q of M and N. Any two lifts φ and ψ of $f\colon M \longrightarrow N$ to $P \longrightarrow Q$ are homotopic, by Theorem 5.19. Additive functors preserve homotopies, by Theorem 7.47, so $F(\varphi)$ and $F(\psi)$ are homotopic. Homotopic maps induce the same map in homology, by Theorem 2.12, so $L_iF(\varphi) = L_iF(\psi)$ for each i.

c) Given an arrow f, fix a lift φ of f to projective resolutions of the source and target, which exists by Theorem 5.19. Since F is an additive functor, $\mathrm{H}_i(F(\varphi))$ is a homomorphism for each i, and thus $L_iF(f)$ is a homomorphism between the corresponding Hom-groups, which as we have seen is independent of our choice of φ.

d) Let A be any R-module and P be a projective resolution of A. Since F is right exact, and

$$P_1 \longrightarrow P_0 \longrightarrow A \longrightarrow 0$$

is exact, then so is

$$F(P_1) \longrightarrow F(P_0) \longrightarrow F(A) \longrightarrow 0.$$

We claim that $\mathrm{H}_0(F(P)) = F(A)$. The last sequence above says that

$$F(A) = \operatorname{coker}(F(P_1) \to F(P_0)),$$

and $\mathrm{H}_0(F(P)) = F(P_0)/\operatorname{im}(F(P_1) \to F(P_0)) = \operatorname{coker}(F(P_1) \to F(P_0))$. □

Exercise 6.5. Show that the following holds for every covariant left exact functor F:

a) $R^iF(A)$ is well-defined up to isomorphism.

b) $R^iF(f)$ is well-defined for every arrow f.

c) $R^iF(f)$ is an additive functor for every i.

d) $R^0F = F$.

And now we are ready to prove the most important result about derived functors: they fix the lack of exactness of the functor we are deriving, by inducing a long exact sequence in homology from any given short exact sequence.

Theorem 6.6. *Let F be a right exact covariant functor. Any short exact sequence*

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

induces a natural long exact sequence

$$\cdots \longrightarrow L_2F(C) \longrightarrow L_1F(A) \longrightarrow L_1F(B) \longrightarrow L_1F(C) \longrightarrow F(A) \longrightarrow F(B) \longrightarrow F(C) \longrightarrow 0.$$

Similarly, if F is a left exact covariant functor, we obtain a long exact sequence

$$0 \longrightarrow F(A) \longrightarrow F(B) \longrightarrow F(C) \longrightarrow R^1F(A) \longrightarrow R^1F(B) \longrightarrow R^1F(C) \longrightarrow R^2F(A) \longrightarrow \cdots .$$

If F is a contravariant left exact functor, we obtain a natural long exact sequence

$$0 \longrightarrow F(C) \longrightarrow F(B) \longrightarrow F(A) \longrightarrow R^1F(C) \longrightarrow R^1F(B) \longrightarrow R^1F(A) \longrightarrow R^2F(C) \longrightarrow \cdots .$$

Proof. We give a proof for the case of right exact functors, and the remaining cases follow by duality. We start by fixing projective resolutions P of A and R of C. By Theorem 5.21, we can choose a projective resolution Q of B and lifts of f and g such that

$$0 \longrightarrow P \longrightarrow Q \longrightarrow R \longrightarrow 0$$

is a short exact sequence of complexes. By Theorem 6.4, L_iF does not depend on the choice of resolution, so we can compute $L_iF(A)$, $L_iF(B)$, and $L_iF(C)$ from P, Q, and R. Now notice that for each n, R_n is projective, so

$$0 \longrightarrow P_n \longrightarrow Q_n \longrightarrow R_n \longrightarrow 0$$

is a split short exact sequence. Now additive functors preserve split short exact sequences, by Theorem 7.72, so

$$0 \longrightarrow F(P_n) \longrightarrow F(Q_n) \longrightarrow F(R_n) \longrightarrow 0$$

is a short exact sequence for all n. Then

$$0 \longrightarrow F(P) \longrightarrow F(Q) \longrightarrow F(R) \longrightarrow 0$$

is a short exact sequence of complexes. Note, however, that this sequence is not necessarily split anymore, since the splittings at each level do not necessarily assemble into a map of complexes. The Long Exact Sequence in homology now gives us the long exact sequence we desire.

There were many choices along the way. First, we chose resolutions P, Q, and R, and lifts of f and g. We have shown our computations of $L_iF(-)$ are independent of these choices. We should check, however, that the resulting connecting arrows are natural transformations that do not depend on our choice of lifts. Once a lift is fixed, we know we already have naturality from the Snake Lemma or the Long Exact Sequence in homology.

It remains to check naturality. What is left to check is that given a commutative diagram with exact rows

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \longrightarrow & B & \longrightarrow & C & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle a} & & \downarrow{\scriptstyle b} & & \downarrow{\scriptstyle c} & & \\ 0 & \longrightarrow & A' & \longrightarrow & B' & \longrightarrow & C' & \longrightarrow & 0 \end{array}$$

and chosen lifts of the original short exact sequences to projective resolutions, there are maps of complexes such that

$$\begin{array}{ccccccccc} 0 & \longrightarrow & P & \longrightarrow & Q & \longrightarrow & R & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \beta} & & \downarrow{\scriptstyle \gamma} & & \\ 0 & \longrightarrow & P' & \longrightarrow & Q' & \longrightarrow & R' & \longrightarrow & 0 \end{array}$$

commutes. Our derived functors L_iF will preserve these maps of complexes and the commutativity of the diagram above, so we get commutative diagrams

$$\begin{array}{ccc} L_iF(C) & \longrightarrow & L_{i-1}F(A) \\ \downarrow{\scriptstyle L_iF(\gamma)} & & \downarrow{\scriptstyle L_{i-1}F(\alpha)} \\ L_iF(C') & \longrightarrow & L_{i-1}F(A) \end{array}$$

for each i. First, notice that we know that a, b, and c can be lifted to maps of complexes by Theorem 5.19, and that any two lifts of each a, b, or c are unique up to homotopy. So let's start by fixing lifts α of a and γ of c, and we will construct an appropriate lift β of b. Since the short exact sequences

$$0 \longrightarrow P_n \longrightarrow Q_n \longrightarrow R_n \longrightarrow 0$$

split for each n, we might as well assume that $Q_n = P_n \oplus R_n$ and that the arrows $P \longrightarrow Q$ and $Q \longrightarrow R$ are given by the canonical arrows to and from the product $\equiv$ coproduct in each homological degree. We cannot, however, assume $Q = P \oplus R$ as complexes, only that $Q_n = P_n \oplus R_n$ in each homological degree n. The commutativity of

$$\begin{array}{ccccc} 0 & \longrightarrow & P_n & \longrightarrow & P_n \oplus R_n \\ & & \downarrow{\scriptstyle \partial_n^P} & & \downarrow{\scriptstyle \partial_n^Q} \\ 0 & \longrightarrow & P_{n-1} & \longrightarrow & P_{n-1} \oplus R_{n-1} \end{array}$$

does imply that $\partial^Q(P) \subseteq P$, so we can say that ∂^Q is of the form

$$\partial_n^Q = \begin{pmatrix} \partial_n^P & \mu_n \\ 0 & \partial_n^R \end{pmatrix}$$

for each n. Since this is a differential, we have

$$(\partial_n^Q)^2 = 0 \implies \partial_{n-1}^P \mu_n + \mu_{n-1}\partial_n^R = 0.$$

Similarly, all this applies to $\partial_n^{Q'}$, which must be of the form

$$\partial_n^{Q'} = \begin{pmatrix} \partial_n^{P'} & \mu'_n \\ 0 & \partial_n^{R'} \end{pmatrix}.$$

We claim that we can define $\beta_n = \begin{pmatrix} \alpha_n & \nu_n \\ 0 & \gamma_n \end{pmatrix}$ for each n such that β is a map of complexes, meaning

$$\partial_n^{Q'}\beta_n = \beta_{n-1}\partial_n^Q.$$

Writing the corresponding products of matrices, we must have

$$\begin{pmatrix} \partial_n^{P'} & \mu'_n \\ 0 & \partial_n^{R'} \end{pmatrix}\begin{pmatrix} \alpha_n & \nu_n \\ 0 & \gamma_n \end{pmatrix} = \begin{pmatrix} \alpha_{n-1} & \nu_{n-1} \\ 0 & \gamma_{n-1} \end{pmatrix}\begin{pmatrix} \partial_n^P & \mu_n \\ 0 & \partial_n^R \end{pmatrix} \implies \begin{cases} \alpha \text{ is a map of complexes} \\ \partial_n^{P'}\nu_n + \mu'_n\gamma_n = \alpha_{n-1}\mu_n + \nu_{n-1}\partial_n^R \\ 0 = 0 \\ \gamma \text{ is a map of complexes} \end{cases}$$

The only nontrivial statement we want to guarantee is that $\partial_n^{P'}\nu_n + \mu'_n\gamma_n = \alpha_{n-1}\mu_n + \nu_{n-1}\partial_n^R$. We can solve this inductively for each n, and construct an appropriate ν_n inductively. Given ν_{n-1}, set

$$\Gamma_n := \alpha_{n-1}\mu_n + \nu_{n-1}\partial_n^R - \mu'_n\gamma_n,$$

We want to construct ν_n such that $\begin{array}{ccc} R_n & \xrightarrow{\nu_n} & P'_n \\ & \searrow^{\Gamma_n} & \downarrow \\ & & P'_{n-1} \end{array}$ commutes, assuming we have constructed ν_{n-1}. First, we claim that $\partial_{n-1}^{P'}\Gamma_n = 0$.

$$\begin{aligned} \partial_{n-1}^{P'}\Gamma_n =& \partial_{n-1}^{P'}\alpha_{n-1}\mu_n + \partial_{n-1}^{P'}\nu_{n-1}\partial_n^R - \partial_{n-1}^{P'}\mu'_n\gamma_n \\ =& \mu'_{n-1}\partial_n^{P'}\gamma_n + \partial_{n-1}^{P'}\alpha_{n-1}\mu_n + \partial_{n-1}^{P'}\nu_{n-1}\partial_n^R \qquad \text{since } \mu'_{n-1}\partial_n^{P'} = \partial_{n-1}^P\mu_n \end{aligned}$$

By induction,

$$\partial_{n-1}^{P'}\nu_{n-1} + \mu'_{n-1}\gamma_{n-1} = \alpha_{n-2}\mu_{n-1} + \nu_{n-2}\partial_{n-1}^R.$$

Using this to replace $\partial_{n-1}^{P'}\nu_{n-1}$ in the equation above, we get

$$\begin{aligned} \partial_{n-1}^{P'}\Gamma_n =& \mu'_{n-1}\partial_n^{P'}\gamma_n + \partial_{n-1}^{P'}\alpha_{n-1}\mu_n + (\alpha_{n-2}\mu_{n-1} + \nu_{n-2}\partial_{n-1}^R - \mu'_{n-1}\gamma_{n-1})\partial_n^R \\ =& \alpha_{n-2}\mu_{n-1}\partial_n^R + \partial_{n-1}^{P'}\alpha_{n-1}\mu_n + \nu_{n-2}\partial_{n-1}^R\partial_n^R - \mu'_{n-1}(\partial_n^{P'}\gamma_n + \gamma_{n-1}\partial_n^R) \\ =& \alpha_{n-2}\partial_{n-1}^P\mu_n + \partial_{n-1}^{P'}\alpha_{n-1}\mu_n + \nu_{n-2}\partial_{n-1}^R\partial_n^R - \mu'_{n-1}(\partial_n^{P'}\gamma_n + \gamma_{n-1}\partial_n^R) \end{aligned}$$

We showed above that $\partial_n^{P'}\gamma_n + \gamma_{n-1}\partial_n^R = 0$. Moreover, $\partial_{n-1}^R\partial_n^R = 0$. We conclude that

$$\begin{aligned}
\partial_{n-1}^{P'}\Gamma_n =& \alpha_{n-2}\partial_{n-1}^P\mu_n + \partial_{n-1}^{P'}\alpha_{n-1}\mu_n \\
=& \alpha_{n-2}\partial_{n-1}^P\mu_n + \alpha_{n-2}\partial_n^{P'}\mu_n && \text{since } \alpha \text{ is a map of complexes} \\
=& \alpha_{n-2}(\partial_{n-1}^P\mu_n + \partial_n^{P'}\mu_n) \\
=& 0 && \text{since } \partial_{n-1}^P\mu_n + \partial_n^{P'}\mu_n = 0.
\end{aligned}$$

So this concludes the proof that $\partial_{n-1}^{P'}\Gamma_n = 0$. Therefore, Γ_n must factor through the $\ker\partial_{n-1}^{P'}$:

$$\begin{array}{ccccc}
P'_n & \xrightarrow{\partial_n} & P'_{n-1} & \xrightarrow{\partial_{n-1}} & P'_{n-2} \\
 & \nwarrow & \uparrow{\scriptstyle \Gamma_n} & & \\
 & \ker\partial_{n-1} & \xleftarrow[\psi_n]{} & R_n &
\end{array}$$

On the other hand, P' is a resolution and thus exact, so $\operatorname{im}\partial_n = \ker\partial_{n-1}$, and ∂_n factors through $\ker\partial_{n-1}$ as

$$\begin{array}{ccccc}
P'_n & \xrightarrow{\partial_n} & P'_{n-1} & \xrightarrow{\partial_{n-1}} & P'_{n-2} \\
 & \searrow^{\varphi_n} \nearrow & \uparrow{\scriptstyle \Gamma_n} & & \\
 & \ker\partial_{n-1} & \xleftarrow[\psi_n]{} & R_n &
\end{array}$$

via some epi φ_n. Finally, R_n is projective, so there exists ν_n such that

$$\begin{array}{ccc}
 & & R_n \\
 & \overset{\nu_n}{\swarrow} & \downarrow{\scriptstyle \psi_n} \\
P'_n & \xrightarrow[\varphi_n]{} & \ker\partial_{n-1}
\end{array}$$

commutes — this was the ν_n we were searching for. □

Theorem 6.7. *Let $T_i : R$-**Mod** $\longrightarrow S$-**Mod** be a sequence of additive covariant functors, and $F : R$-**Mod** $\longrightarrow S$-**Mod** a right exact functor. Suppose that the following hold:*

*(1) For every short exact sequence $0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$ in R-**Mod**, we get a natural long exact sequence*

$$\cdots \longrightarrow T_2(C) \longrightarrow T_1(A) \longrightarrow T_1(B) \longrightarrow T_1(C) \longrightarrow T_0(A) \longrightarrow T_0(B) \longrightarrow T_0(C) \longrightarrow 0.$$

(2) T_0 is naturally isomorphic to F.

(3) $T_n(P) = 0$ for every projective P and all $n \geqslant 1$.

Then T_n is naturally isomorphic to L_nF for all $n \geqslant 0$.

Proof. We are going to show that T_n is naturally isomorphic to L_nF by all n. The statement for $n = 0$ is one of our assumptions. When $n = 1$, fix an R-module M, and consider a short exact sequence

$$0 \longrightarrow K \xrightarrow{f} P \longrightarrow M \longrightarrow 0$$

with P projective. By assumption (1), we get a long exact sequence on the T_i, and by (2), there exist isomorphisms τ_0 such that the following is a commutative diagram:

$$\begin{array}{ccccccccc} T_1(P) & \longrightarrow & T_1(M) & \xrightarrow{\Delta_1} & T_0(K) & \xrightarrow{T_0(f)} & T_0(P) & \longrightarrow & T_0(M) \longrightarrow 0 \\ & & & & \downarrow{\scriptstyle \tau_0(K)} & & \downarrow{\scriptstyle \tau_0(P)} & & \downarrow{\scriptstyle \tau_0(M)} \\ L_1F(P) & \longrightarrow & L_1F(M) & \xrightarrow[\delta_1]{} & F(K) & \xrightarrow[F(f)]{} & F(P) & \longrightarrow & F(M) \longrightarrow 0. \end{array}$$

By (3), $T_1(P) = 0$, and $L_1F(P) = 0$ by construction. The exactness of each row now implies that Δ_1 and δ_1 are both injective. Moreover,

$$\begin{aligned} F(f)\tau_0(K)\Delta_1 &= \tau_0(P)T_0(f)\Delta_1 && \text{by commutativity of the diagram} \\ &= 0 && \text{since } T_0(f)\Delta_1 = 0. \end{aligned}$$

so the image of $\tau_(k)\Delta_1$ is contained in $\ker F(f) = \operatorname{im}\delta_1$. Define $\tau_1(M)\colon T_1(M) \to L_1F(M)$ as follows: we send each $a \in T_1(M)$ to the unique $b \in L_1F(M)$ such that $\delta_1(b) = \tau_0(K)\Delta_1(a)$. This is a homomorphism of R-modules because so are δ_1, $\tau_0(K)$, and Δ_1. Moreover, since $\tau_0(K)$ is an isomorphism and Δ_1 is injective, the composition $\tau_0(K)\Delta_1$ is injective. As a consequence, $\tau_1(M)$ is injective. On the other hand, we claim that $\tau_1(M)$ is also surjective. Given any $b \in L_1F(M)$, since $\tau_0(K)$ is an isomorphism there exists $c \in T_0(K)$ such that $\tau_0(K)(c) = \delta_1(b)$. Thus

$$\begin{aligned} \tau_0(P)T_0(f)(c) &= F(f)\tau_0(K)(c) && \text{by commutativity} \\ &= F(f)\delta_1(b) && \text{since } \tau_0(K)(c) = \delta_1(b) \\ &= 0 && \text{since the bottom row is a complex} \end{aligned}$$

Since $\tau_0(P)$ is an iso, we must have $c \in \ker(T_0(f)) = \operatorname{im}\Delta_1$. Thus we can choose $a \in T_1(M)$ such that $\Delta_1(a) = c$, which implies that $\tau_1(M)(a) = b$. Therefore, $\tau_1(M)$ is an isomorphism.

This shows that $T_1(M) \cong L_1F(M)$. Now let $n \geqslant 1$, and consider the diagram with exact rows

$$\begin{array}{ccccccc} T_{n+1}(P) & \longrightarrow & T_{n+1}(M) & \xrightarrow{\Delta_{n+1}} & T_n(K) & \longrightarrow & T_n(P) \\ & & & & \downarrow{\scriptstyle \tau_n(K)} & & \\ L_{n+1}F(P) & \longrightarrow & L_{n+1}F(M) & \xrightarrow[\delta_{n+1}]{} & L_nF(K) & \longrightarrow & L_nF(P) \end{array}$$

By (3), $T_{n+1}(P) = 0 = T_n(P)$, and by construction $L_{n+1}F(P) = 0 = L_nF(P)$. Therefore, Δ_{n+1} and δ_{n+1} are isomorphisms. Since $\tau_n(K)$ is also an isomorphism, we conclude that $T_{n+1}(M) \cong L_{n+1}F(M)$. Therefore, $T_n(M) \cong L_nF(M)$ for all n.

It remains to show that these isomorphisms are natural, that is, that any R-module map $f\colon M \to N$ gives rise to commutative diagrams

$$\begin{array}{ccc} T_i(M) & \xrightarrow{\tau_i(M)} & L_iF(M) \\ {\scriptstyle T_i(f)}\downarrow & & \downarrow{\scriptstyle L_iF(f)} \\ T_i(N) & \xrightarrow[\tau_i(N)]{} & L_iF(N). \end{array}$$

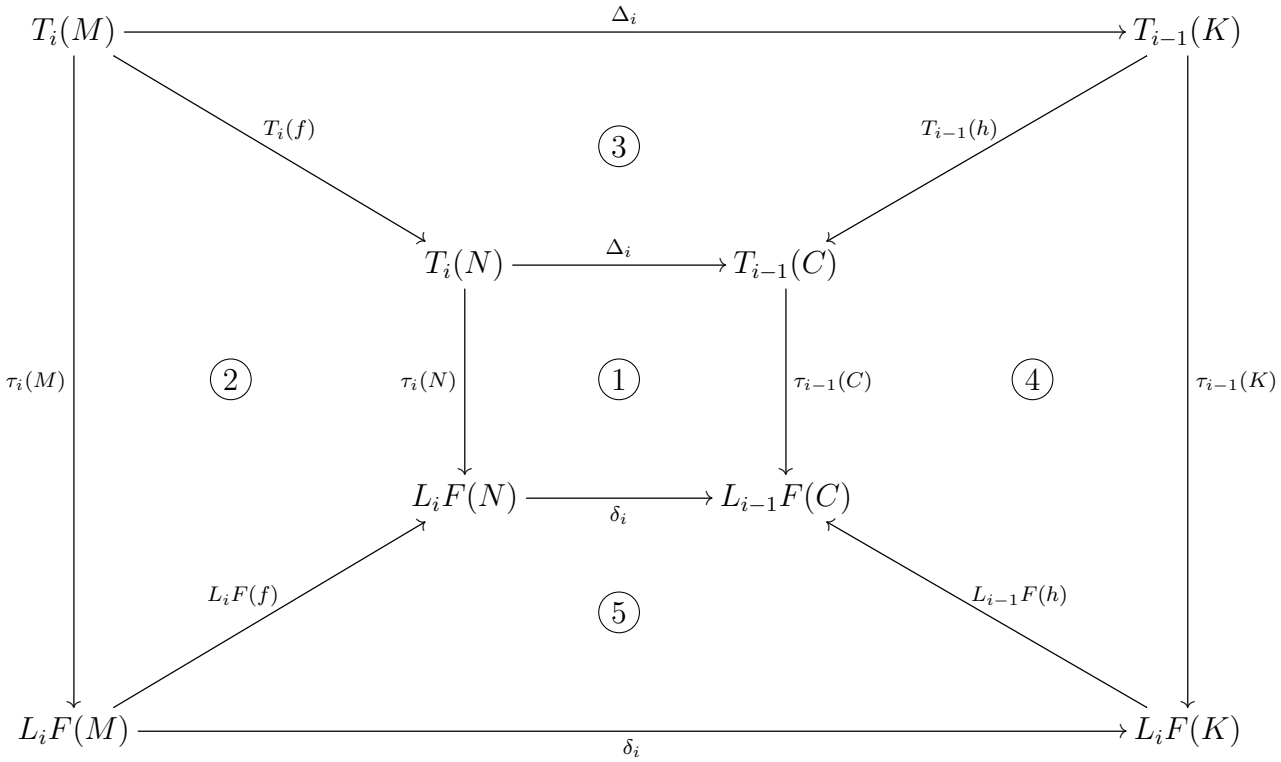
We will prove this by induction on i. First, note that the commutativity of the square holds for $i = 0$ by (2). Let $i \geqslant 1$. Fix projectives P and Q and short exact sequences

$$0 \longrightarrow K \longrightarrow P \longrightarrow M \longrightarrow 0 \qquad \text{and} \qquad 0 \longrightarrow C \longrightarrow Q \longrightarrow N \longrightarrow 0.$$

Since Q is projective and $Q \to N$ is surjective, f lifts to a map $g : P \to Q$. Moreover, an argument similar to the one we used above shows that we can define a map h giving a commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & K & \longrightarrow & P & \longrightarrow & M & \longrightarrow & 0 \\ & & {\scriptstyle h}\downarrow & & {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle f} & & \\ 0 & \longrightarrow & C & \longrightarrow & Q & \longrightarrow & N & \longrightarrow & 0. \end{array}$$

Now consider the following diagram:



- The big square and (1) commute by definition of τ_i.
- The square (3) commutes because we assumed in (1) that T_i gives rise to long exact sequences which are natural.
- The square (5) commutes because L_iF gives rise to natural long exact sequences, by Theorem 6.6.
- The square (4) commutes by induction hypothesis.

Our goal is to show that (2) commutes. First, we claim that

$$\delta_i \circ \tau_i(N) \circ T_i(f) = \delta_i \circ L_iF(f) \circ \tau_i(M).$$

Indeed, using the commutativity of the various other parts of the diagram, we get

$$\begin{aligned}
\delta_i \circ \tau_i(N) \circ T_i(f) &= \tau_{i-1}(C) \circ \Delta_i \circ T_i(f) && \text{by commutativity of (1)} \\
&= \tau_{i-1}(C) \circ T_{i-1}(h) \circ \Delta_i && \text{by commutativity of (3)} \\
&= L_{i-1}F(h) \circ \tau_{i-1}(K) \circ \Delta_i && \text{by commutativity of (4)} \\
&= L_{i-1}F(h) \circ \delta_i \circ \tau_i(M) && \text{by commutativity of the big square} \\
&= \delta_i \circ L_iF(f) \circ \tau_i(M) && \text{by commutativity of (5).}
\end{aligned}$$

On the other hand, the long exact sequence for L_iF from Theorem 6.6 says that

$$L_iF(Q) \longrightarrow L_iF(N) \xrightarrow{\delta_i} L_{i-1}F(C)$$

is exact, but since $i \geqslant 1$ and Q is projective we have $L_iF(Q) = 0$ by Theorem 6.3. But the exactness of

$$0 \longrightarrow L_iF(N) \xrightarrow{\delta_i} L_{i-1}F(C)$$

says that δ_i is injective. Therefore,

$$\delta_i \circ \tau_i(N) \circ T_i(f) = \delta_i \circ L_iF(f) \circ \tau_i(M) \implies \tau_i(N) \circ T_i(f) = L_iF(f) \circ \tau_i(M),$$

and (2) commutes, as desired. □

There are versions of this theorem for the three remaining cases as well; we record one of them here:

Theorem 6.8. *Suppose $T_i \colon R\text{-}\boldsymbol{Mod} \longrightarrow S\text{-}\boldsymbol{Mod}$ is a sequence of additive covariant functors and $F \colon R\text{-}\boldsymbol{Mod} \longrightarrow S\text{-}\boldsymbol{Mod}$ a left exact functor such that*

*a) For every short exact sequence $0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$ in R-**Mod**, we get a long exact sequence*

$$0 \longrightarrow T_0(A) \longrightarrow T_0(B) \longrightarrow T_0(C) \longrightarrow T_1(A) \longrightarrow T_1(B) \longrightarrow T_1(C) \longrightarrow \cdots.$$

b) T_0 is naturally isomorphic to F.

c) $T_n(E) = 0$ for every injective E and all $n \geqslant 1$.

Then T_n is naturally isomorphic to R^nF for all n.

We leave the proof of this and the other two cases as an exercise.

6.2 A first look at Ext and Tor

It's time to study some concrete examples of derived functors: Ext, the derived functor of Hom, and Tor, the derived functor of tensor. Given two modules M and N, we may consider the derived functors of $M \otimes_R -$, and then plug in N, or we may consider the derived functors of $- \otimes_R N$, and plug in M; it turns out that the two are naturally isomorphic, and this is the Tor functor:

$$\mathrm{Tor}_i^R(M,N) := L_i(M \otimes_R -)(N) \cong L_i(- \otimes_R N)(M).$$

More precisely, if P is a projective resolution of M, and Q is a projective resolution of N,

$$\mathrm{Tor}_i^R(M,N) := \mathrm{H}_i(P \otimes_R N) \cong \mathrm{H}_i(M \otimes_R Q).$$

There are two Hom functors, each with its own derived functor: given R-modules M and N, we may take a projective resolution P of M, and compute $\mathrm{H}^i(\mathrm{Hom}_R(P,N))$, or we could take an injective resolution E of N, and compute $\mathrm{H}^i(\mathrm{Hom}_R(M,E))$. It turns out these two completely different sounding constructions give us isomorphic R-modules:

$$\begin{aligned}\mathrm{Ext}_R^i(M,N) &:= R^i(\mathrm{Hom}_R(M,-))(N) \cong R^i(\mathrm{Hom}_R(-,N))(M) \\ &\cong \mathrm{H}^i(\mathrm{Hom}_R(P,N)) \cong \mathrm{H}^i(\mathrm{Hom}_R(M,E)).\end{aligned}$$

To show that for each of Ext and Tor these two seemingly unrelated definitions agree, we will need some more tools.

Definition 6.9. The **suspension** of **shift** of a complex C is the complex $\Sigma C := C[1]$ with

$$(\Sigma C)_n = C_{n+1} \quad \text{and} \quad \partial^{\Sigma C} = -\partial^C.$$

Given an integer k, the k**th suspension** of C is the complex

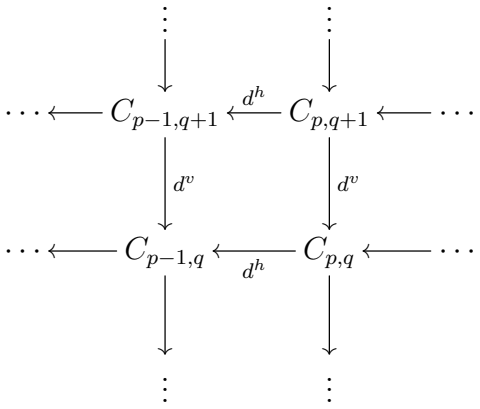
$$\Sigma^k C := \underbrace{\Sigma \cdots \Sigma}_{k \text{ times}} C \qquad \text{with } \partial^{\Sigma^k C} = (-1)^k \partial^C.$$

In particular, $\Sigma^{-1}C$ is the complex with

$$(\Sigma C)_n = C_{n-1} \quad \text{and} \quad \partial^{\Sigma^{-1} C} = -\partial^C.$$

Note that there are two conventions in the literature, the other one being $(\Sigma C)_n = C_{n-1}$.

Definition 6.10. A (homological) **double complex** over the ring R is a family of R-modules $\{C_{p,q}\}_{p,q \in \mathbb{Z}}$ together with homomorphisms of R-modules $d^h : C_{p,q} \longrightarrow C_{p-1,q}$ and $d^v : C_{p,q} \longrightarrow C_{p,q-1}$



satisfying

$$d^h d^h = 0 \qquad d^v d^v = 0 \qquad d^h d^v + d^v d^h = 0.$$

Remark 6.11. Note that if C is a double complex, then each row and each column is a complex: if we fix p, $C_{p,\bullet}$ is a complex with differential d^v; if we fix q, $C_{\bullet,q}$ is a complex with differential d^h.

What we defined above is a homological double complex. A cohomological double complex would have vertical and horizontal maps that go up in index, and we instead write $C^{p,q}$ for the module in position (p, q). Also, please note that there are different conventions in the literature for whether p refers to the row or column.

Definition 6.12. Given a double complex C, its **total complex** is the complex given by

$$\mathrm{Tot}^{\oplus}(C)_n := \bigoplus_{p+q=n} C_{p,q} \quad \text{with differential } d = d^h + d^v.$$

Similarly, the **product total complex** of C is given by

$$\mathrm{Tot}^{\prod}(C)_n := \prod_{p+q=n} C_{p,q} \quad \text{with differential } d = d^h + d^v.$$

Remark 6.13. Let C be a double complex with differentials d^v and d^h. Then

$$(d^h + d^v)(d^h + d^v) = \underbrace{d^h d^h}_{0} + \underbrace{d^h d^v + d^v d^h}_{0} + \underbrace{d^v d^v}_{0} = 0,$$

so $(\mathrm{Tot}^{\oplus}(C), d)$ and $(\mathrm{Tot}^{\prod}(C), d)$ are indeed complexes.

In order to prove our two definitions of Ext and Tor each agree, we will need two special double complexes: the tensor and the Hom double complex.

Definition 6.14. Let R be a ring and C and D be complexes of R-modules. The **tensor product double complex** of C and D is the double complex $C \otimes D$ given by taking

$$(C \otimes D)_{p,q} = C_p \otimes D_q \quad d^h = \partial^C \otimes_R 1_D, \quad \text{and} \quad d^v = (-1)^p 1_C \otimes_R \partial^D.$$

We call the total complex of the tensor product double complex of C and D the **tensor product** of C and D in $\mathrm{Ch}(R)$, and denote it by $C \otimes D$.

Remark 6.15. The tensor product total complex has

$$\mathrm{Tot}^{\oplus}(C \otimes D)_n = \bigoplus_{p+q=n} C_p \otimes_R D_q$$

and differential

$$d(x \otimes y) = \partial(x) \otimes y + (-1)^p x \otimes \partial(y)$$

for $x \in C_p$ and $y \in D_q$.

Definition 6.16. Let R be a ring and C and D be complexes of R-modules. The **Hom double complex** of C and D is the double complex $\mathrm{Hom}(C, D)$ given by

$$(\mathrm{Hom}(C, D))_{p,q} := \mathrm{Hom}_R(C_{-p}, D_q)$$

with differentials

$$\begin{array}{ccc} \mathrm{Hom}_R(C_{-p}, D_q) & \xrightarrow{d^h} & \mathrm{Hom}_R(C_{-p+1}, D_q) \\ f & \longmapsto & f \circ \partial^C \end{array} \quad \text{and} \quad \begin{array}{ccc} \mathrm{Hom}_R(C_{-p}, D_q) & \xrightarrow{d^v} & \mathrm{Hom}_R(C_{-p}, D_{q-1}) \\ f & \longmapsto & (-1)^{p+q+1} \partial^D \circ f \end{array}.$$

We call the product total complex of the Hom double complex of C and D the **(internal) Hom complex** of C and D, and denote it by $\mathrm{Hom}(C, D)$.

Remark 6.17. The Hom complex of C and D is the complex

$$\mathrm{Hom}(C, D)_n = \prod_{p+q=n} \mathrm{Hom}_R(C_{-p}, D_q)$$

with differential $d(f) = f \circ \partial^C + (-1)^{p+q+1} \partial^D \circ f$ for each $f \in \mathrm{Hom}_R(C_{-p}, D_q)$.

Remark 6.18. Given C and D in $\mathrm{Ch}(R)$, what is a 0-cycle in the Hom complex $\mathrm{Hom}(C, D)$? A 0-cycle is a sequence of maps of R-modules $f_k \colon C_k \to D_k$ satisfying $f\partial^C - \partial^D f = 0$, so the 0-cycles are precisely the maps of complexes $C \to D$. Similarly, a sequence of maps $f_k \colon C_k \to D_k$ is a 0-boundary if there exists a sequence of maps $h_k \colon C_k \longrightarrow D_{k+1}$ such that $f_k = \partial^D h_k + h_{k-1} \partial^C$. In other words, a 0-boundary indicates a homotopy relation — if $f - g$ is a 0-boundary, f and g are homotopic maps.

Definition 6.19. Let $f \colon C \longrightarrow D$ be a map of complexes. The **(mapping) cone** of f is the complex $\mathrm{cone}(f)$ with $\mathrm{cone}(f)_n = C_{n-1} \oplus D_n$ and differential given by

$$\partial_n := \begin{pmatrix} -\partial_C & 0 \\ f & \partial_D \end{pmatrix} : \begin{array}{ccc} C_{n-1} & \xrightarrow{-\partial^C} & C_{n-2} \\ \oplus & \searrow^{f} & \oplus \\ D_n & \xrightarrow[\partial^D]{} & D_{n-1} \end{array}$$

Remark 6.20. There are different conventions for the sign in front of f in the definition of the differentials on the cone of f. Weibel [Wei94] defines

$$\partial_n := \begin{pmatrix} -\partial_C & 0 \\ -f & \partial_D \end{pmatrix}.$$

and some authors even write

$$\partial_n := \begin{pmatrix} \partial_C & 0 \\ (-1)^{n-1} f & \partial_D \end{pmatrix}.$$

Both of these choices do make our proposed differential a differential (check it!). The facts below about the mapping cone are all true up to sign whatever the sign convention we follow.

Exercise 6.21. Let $f \colon A \to B$ be a map of complexes. Show that f is nullhomotopic if and only if f factors through the canonical map $A \to \mathrm{cone}(\mathrm{id}_A)$.

Exercise 6.22. Let $f\colon C \to D$ be a map of complexes. Show that giving a map of complexes $\mathrm{cone}(f) \to E$ is the same as giving

- a map of complexes $D \xrightarrow{g} E$, and
- a homotopy between gf and 0.

Remark 6.23. Given any map of complexes $C \xrightarrow{f} D$, there is a short exact sequence

$$0 \longrightarrow D \longrightarrow \mathrm{cone}(f) \longrightarrow \Sigma^{-1}C \longrightarrow 0$$

determined by the canonical arrows to and from the product $\equiv$ coproduct. The connecting arrows from the Snake Lemma

$$\mathrm{H}_{n-1}(C) = \mathrm{H}_n(\Sigma^{-1}C) \xrightarrow{\delta} \mathrm{H}_{n-1}(D)$$

are exactly $\mathrm{H}_{n-1}(f)\colon \mathrm{H}_{n-1}(C) \longrightarrow \mathrm{H}_{n-1}(D)$ induced by f, so there is a long exact sequence

$$\cdots \longrightarrow \mathrm{H}_{n+1}(\mathrm{cone}(f)) \longrightarrow \mathrm{H}_n(C) \xrightarrow{\mathrm{H}_n(f)} \mathrm{H}_n(D) \longrightarrow \mathrm{H}_n(\mathrm{cone}(f)) \longrightarrow \mathrm{H}_{n-1}(C) \longrightarrow \cdots.$$

As a consequence, f is a quasi-isomorphism if and only if $\mathrm{cone}(f)$ is exact.

Remark 6.24. Given a map of complexes $C \xrightarrow{f} D$, we can construct a double complex from f, as follows:

$$X = \begin{array}{ccccccc}
 & & \vdots & & \vdots & & \\
 & & \downarrow{\scriptstyle \partial} & & \downarrow{\scriptstyle -\partial} & & \\
0 & \longleftarrow & D_2 & \xleftarrow{f} & C_2 & \longleftarrow & 0 \\
 & & \downarrow{\scriptstyle \partial} & & \downarrow{\scriptstyle -\partial} & & \\
0 & \longleftarrow & D_1 & \xleftarrow{f} & C_1 & \longleftarrow & 0 \\
 & & \downarrow{\scriptstyle \partial} & & \downarrow{\scriptstyle -\partial} & & \\
0 & \longleftarrow & D_0 & \xleftarrow{f} & C_0 & \longleftarrow & 0 \\
 & & \downarrow{\scriptstyle \partial} & & \downarrow{\scriptstyle -\partial} & & \\
 & & \vdots & & \vdots & &
\end{array}$$

Note that $\mathrm{Tot}^{\oplus}(X) = \mathrm{cone}(f)$.

Now that we have introduced all the tools we need, the last thing we need is a technical but very useful lemma.

Lemma 6.25 (Acyclic Assembly Lemma)**.** *Let C be a double complex in R-**Mod**.*

a) If C is an upper half plane double complex with exact rows, meaning $C_{p,q} = 0$ whenever $q < 0$, then $\mathrm{Tot}^{\oplus}(C)$ is exact.

b) If C is a right half plane double complex with exact columns, meaning $C_{p,q} = 0$ whenever $p < 0$, then $\mathrm{Tot}^{\oplus}(C)$ is exact.

c) If C is an upper half plane double complex with exact columns, meaning $C_{p,q} = 0$ whenever $q < 0$, then $\mathrm{Tot}^{\prod}(C)$ is exact.

d) If C is a right half plane double complex with exact rows, meaning $C_{p,q} = 0$ whenever $p < 0$, then $\mathrm{Tot}^{\prod}(C)$ is exact.

Proof. Notice that a) $\Leftrightarrow$ b) and c) $\Leftrightarrow$ d) by switching the indexes. Moreover, we claim that it is sufficient to show c), since it implies b).

To show that c) implies b), we need some notation. Given a double complex C, consider the nth truncation $\tau_n(C)$ of C defined by

$$\tau_n(C)_{p,q} := \begin{cases} C_{p,q} & \text{if } q > n \\ \ker(\, C_{p,n} \xrightarrow{d^v} C_{p,n-1}\,) & \text{if } q = n \\ 0 & \text{if } q < n. \end{cases}$$

Suppose that C is a right half plane double complex with exact columns, and assume that c) holds. Then $\tau_n(C)$ still has exact columns, so by c), $\mathrm{Tot}^{\prod}(\tau_n(C))$ is exact. On the other hand, notice that up to a vertical shift, $\tau_n(C)$ is a first quadrant double complex, and for each fixed m, there are only finitely many values of p and q with $p + q = m$ and such that $\tau_n(C)_{p,q} \neq 0$. Therefore, $\mathrm{Tot}^{\prod}(\tau_n(C_{p,\bullet})) = \mathrm{Tot}^{\oplus}(\tau_n(C_{p,\bullet}))$, so $\mathrm{Tot}^{\oplus}(\tau_n(C_{p,\bullet}))$ is exact. We claim that this implies that $\mathrm{Tot}^{\oplus}(C)$ is exact. One can make this precise by saying $\mathrm{Tot}^{\oplus}(C) = \mathrm{colim}_n(\mathrm{Tot}^{\oplus}(C))$. The point is that any element $a \in Z_k(\mathrm{Tot}^{\oplus}(C))$, when we write a explicitly as $a = (a_{p,q}) \in \oplus_{p+q=k} C_{p,q}$ in terms of its coordinates in each $C_{p,q}$, only finitely many $a_{p,q}$ are nonzero. Let q be the smallest such that $a_{p,q} \neq 0$, and fix any $n < q$. Then

$$a \in Z_k(\mathrm{Tot}^{\oplus}(\tau_n(C))) = B_k(\mathrm{Tot}^{\oplus}(\tau_n(C))) \subseteq B_k(\mathrm{Tot}^{\oplus}(C)).$$

So $\mathrm{Tot}^{\oplus}(C)$ is exact, and b) holds.

All we have left to do is to show c), meaning that the product total complex of any upper half plane double complex C with exact columns is exact. We are going to show that $\mathrm{H}_0(\mathrm{Tot}^{\prod}(C)) = 0$, and the remaining homologies follow by shifting C left and right. Consider a 0-cycle in $\mathrm{Tot}^{\prod}(C)$, meaning a sequence of elements $c_p \in C_{-p,p}$ for each $p \geqslant 0$ such that $c = (c_p) \in Z_0(\mathrm{Tot}^{\prod}(C))$. So

$$d(c) = 0 \Leftrightarrow d^v(c_p) + d^h(c_{p-1}) = 0 \text{ for all } p.$$

We will construct $b_{-p,p+1} \in C_{-p,p+1}$ for each p such that $d^v(b_{-p,p+1}) + d^h(b_{-p+1,p}) = c_p$, proving that $c \in B_0(\mathrm{Tot}^{\prod}(C))$.

Set $b_{1,0} = 0 \in C_{1,0}$ when $p = -1$. Since $C_{0,-1} = 0$, we must have $d^v(c_0) = 0 \in C_{0,-1}$. We also assumed that the columns are exact, so in particular the 0th column is exact. We can then find $b_{0,1} \in C_{0,1}$ such that $d^v(b_{0,1}) = c_0$, and thus $d^v(b_{0,1}) + d^h(b_{1,0}) = c_0$.

Now we proceed by induction. Suppose we have constructed $b_{-s+1,s}$ for $-1 \leqslant s \leqslant p$ with the desired property that $d^v(b_{-s,s+1}) + d^h(b_{-s+1,s}) = c_s$ for all $s \leqslant p$. Then

$$\begin{aligned}
d^v(c_{-p,p} - d^h(b_{-p+1,p})) &= d^v(c_p) + d^h d^v(b_{-p+1,p}) && \text{since } d^v d^h + d^h d^v = 0 \\
&= d^v(c_p) + d^h(c_{p-1} - d^h(b_{-p+2,p-1})) && \text{as } d^v(b_{-p+1,p}) + d^h(b_{-p+2,p-1}) = c_{p-1} \\
&= d^v(c_p) + d^h(c_{p-1}) - d^h d^h(b_{-p+2,p-1}) \\
&= d^v(c_p) + d^h(c_{p-1}) && \text{since } d^h d^h = 0 \\
&= 0.
\end{aligned}$$

The last equality comes simply from the fact that $(d^v + d^h)(c) = 0$. So we have shown that $d^v(c_{-p,p} - d^h(b_{-p+1,p})) = 0$. Since the columns are exact, we can find $b_{-p,p+1} \in C_{-p,p+1}$ such that

$$d^v(b_{-p,p+1}) = c_{-p,p} - d^h(b_{-p+1,p}).$$

Equivalently,

$$d^v(b_{-p,p+1}) + d^h(b_{-p+1,p}) = c_{-p,p}. \qquad \square$$

Exercise 6.26. Given a double complex C with $C_{p,q} = 0$ for all $p < n$, the horizontal differentials $C_{n+1,q} \longrightarrow C_{n,q}$ induce a map of complexes

$$\operatorname{Tot}^{\oplus}(C_{>n,\bullet}) \xrightarrow{\varphi} C_{n,\bullet}\,,$$

where $C_{>n,\bullet}$ denotes the double complex we obtain from C by excluding the leftmost nonzero column, and $\operatorname{Tot}^{\oplus}(C) \cong \Sigma^{-1}\operatorname{cone}(\varphi)$, or equivalently, $\Sigma\operatorname{Tot}^{\oplus}(C) \cong \operatorname{cone}(\varphi)$.

We are finally ready to show that the two definitions of Tor coincide.

Theorem 6.27 (Balancing Tor)**.** *Let M and N be R-modules, and fix projective resolutions P of M and Q of N. For all $n \geqslant 0$, there is an isomorphism*

$$L_n(M \otimes_R -)(N) = \mathrm{H}_n(M \otimes_R Q) \cong \mathrm{H}_n(P \otimes_R N) = L_n(- \otimes_R N)(M).$$

Proof. Consider $\pi\colon P_0 \twoheadrightarrow M$ and $\varepsilon\colon Q_0 \twoheadrightarrow N$ and the first quadrant double complex

$$P \otimes Q = \quad \begin{array}{ccccccc}
\vdots & & \vdots & & \vdots & & \\
\downarrow & & \downarrow & & \downarrow & & \\
P_0 \otimes Q_2 & \xleftarrow{\partial^P \otimes 1} & P_2 \otimes Q_1 & \xleftarrow{\partial^P \otimes 1} & P_2 \otimes Q_2 & \longleftarrow & \cdots \\
\downarrow{\scriptstyle 1\otimes\partial^Q} & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & \\
P_0 \otimes Q_1 & \xleftarrow{\partial^P \otimes 1} & P_1 \otimes Q_1 & \xleftarrow{\partial^P \otimes 1} & P_2 \otimes Q_1 & \longleftarrow & \cdots \\
\downarrow{\scriptstyle 1\otimes\partial^Q} & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & \\
P_0 \otimes Q_0 & \xleftarrow{\partial^P \otimes 1} & P_1 \otimes Q_0 & \xleftarrow{\partial^P \otimes 1} & P_2 \otimes Q_0 & \longleftarrow & \cdots
\end{array}$$

Each P_i and Q_i is projective and thus flat, by Theorem 4.39, so $P_i \otimes_R -$ and $- \otimes_R Q_i$ are both exact functors. The rows and columns of our double complex are thus exact everywhere except for the 0th row and column. We can complete our double complex to make a double complex C with both exact rows if we add in a column induced by the surjection π:

$$C = \begin{array}{ccccccccc}
\vdots & & \vdots & & \vdots & & \vdots & & \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
M \otimes Q_2 & \xleftarrow[\pi\otimes 1]{} & P_0 \otimes Q_2 & \xleftarrow{\partial^P \otimes 1} & P_2 \otimes Q_1 & \xleftarrow{\partial^P \otimes 1} & P_2 \otimes Q_2 & \longleftarrow & \cdots \\
\downarrow & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & \\
M \otimes Q_1 & \xleftarrow[\pi\otimes 1]{} & P_0 \otimes Q_1 & \xleftarrow{\partial^P \otimes 1} & P_1 \otimes Q_1 & \xleftarrow{\partial^P \otimes 1} & P_2 \otimes Q_1 & \longleftarrow & \cdots \\
\downarrow & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & \\
M \otimes Q_0 & \xleftarrow[\pi\otimes 1]{} & P_0 \otimes Q_0 & \xleftarrow{\partial^P \otimes 1} & P_1 \otimes Q_0 & \xleftarrow{\partial^P \otimes 1} & P_2 \otimes Q_0 & \longleftarrow & \cdots .
\end{array}$$

We can also make a double complex D with exact columns by adding in a row induced by ε:

$$D = \begin{array}{ccccccccc}
\vdots & & \vdots & & \vdots & & & & \\
\downarrow & & \downarrow & & \downarrow & & & & \\
P_0 \otimes Q_2 & \xleftarrow{\partial^P \otimes 1} & P_2 \otimes Q_1 & \xleftarrow{\partial^P \otimes 1} & P_2 \otimes Q_2 & \longleftarrow & \cdots & & \\
\downarrow{\scriptstyle 1\otimes\partial^Q} & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & & & \\
P_0 \otimes Q_1 & \xleftarrow{\partial^P \otimes 1} & P_1 \otimes Q_1 & \xleftarrow{\partial^P \otimes 1} & P_2 \otimes Q_1 & \longleftarrow & \cdots & & \\
\downarrow{\scriptstyle 1\otimes\partial^Q} & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & \downarrow{\scriptstyle 1\otimes\partial^Q} & & & & \\
P_0 \otimes Q_0 & \xleftarrow{\partial^P \otimes 1} & P_1 \otimes Q_0 & \xleftarrow{\partial^P \otimes 1} & P_2 \otimes Q_0 & \longleftarrow & \cdots & & \\
\downarrow{\scriptstyle 1\otimes\varepsilon} & & \downarrow{\scriptstyle 1\otimes\varepsilon} & & \downarrow{\scriptstyle 1\otimes\varepsilon} & & & & \\
P_0 \otimes N & \longleftarrow & P_1 \otimes N & \longleftarrow & P_2 \otimes N & \longleftarrow & \cdots & & .
\end{array}$$

By Theorem 6.25, $\mathrm{Tot}^{\oplus}(C)$ and $\mathrm{Tot}^{\oplus}(D)$ are both exact. Notice that $\pi \otimes Q$ is a map of complexes $\mathrm{Tot}^{\oplus}(P \otimes Q) \longrightarrow M \otimes Q$, and $P \otimes \varepsilon$ is a map of complexes $\mathrm{Tot}^{\oplus}(P \otimes Q) \longrightarrow P \otimes N$. By Theorem 6.26,

$$\mathrm{cone}(\pi \otimes Q) = \Sigma\,\mathrm{Tot}^{\oplus}(C) \qquad \text{and} \qquad \mathrm{cone}(P \otimes \varepsilon) = \Sigma\,\mathrm{Tot}^{\oplus}(D).$$

Since $\Sigma\,\mathrm{Tot}^{\oplus}(C)$ and $\Sigma\,\mathrm{Tot}^{\oplus}(D)$ are both exact, by Theorem 6.23 both

$$\mathrm{Tot}^{\oplus}(P \otimes Q) \xrightarrow{\pi\otimes Q} M \otimes Q \qquad \text{and} \qquad \mathrm{Tot}^{\oplus}(P \otimes Q) \xrightarrow{P\otimes\varepsilon} P \otimes N$$

are quasi-isomorphisms, so that

$$L_n(M \otimes_R -)(N) = \mathrm{H}_n(M \otimes_R Q) \cong \mathrm{H}_n(P \otimes_R N) = L_n(- \otimes_R N)(M). \qquad \square$$

Theorem 6.28 (Balancing Ext)**.** *Let M and N be R-modules, and fix a projective resolution P of M and an injective resolution E of N. For all n, there is an isomorphism*

$$R^n \operatorname{Hom}_R(M,-)(N) = \mathrm{H}^n(\operatorname{Hom}_R(M,E)) \cong \mathrm{H}^n(\operatorname{Hom}_R(P,N)) = R^n \operatorname{Hom}_R(-,N)(M).$$

Proof. We have a surjection $\pi : P_0 \longrightarrow M$ and an inclusion $\varepsilon : M \longrightarrow E_0$. The double cocomplex $\operatorname{Hom}_R(P,E)$ with $\operatorname{Hom}_R(P,E)_{p,q} = \operatorname{Hom}_R(P_p, E^q)$ and

$$\begin{array}{ccc} \operatorname{Hom}_R(P_p,E^q) \xrightarrow{d^h} \operatorname{Hom}_R(P_{p+1},E^q) & \text{and} & \operatorname{Hom}_R(P_p,E^q) \xrightarrow{d^v} \operatorname{Hom}_R(P_p,D_{q+1}) \,. \\ f \longmapsto f \circ \partial^P & & f \longmapsto (-1)^{p+q+1} \partial^E \circ f \end{array}$$

is a cohomological first quadrant double complex:

$$\operatorname{Hom}(P,E) = \begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \\ \uparrow & & \uparrow & & \uparrow & & \\ \operatorname{Hom}_R(P_0,E^2) & \longrightarrow & \operatorname{Hom}_R(P_1,E^2) & \longrightarrow & \operatorname{Hom}_R(P_2,E^2) & \longrightarrow & \cdots \\ \uparrow & & \uparrow & & \uparrow & & \\ \operatorname{Hom}_R(P_0,E^1) & \longrightarrow & \operatorname{Hom}_R(P_1,E^1) & \longrightarrow & \operatorname{Hom}_R(P_2,E^1) & \longrightarrow & \cdots \\ \uparrow & & \uparrow & & \uparrow & & \\ \operatorname{Hom}_R(P_0,E^1) & \longrightarrow & \operatorname{Hom}_R(P_1,E^1) & \longrightarrow & \operatorname{Hom}_R(P_2,E^1) & \longrightarrow & \cdots \end{array}$$

We proceed just like in Theorem 6.27, now considering two cohomological double complexes:

$$C = \begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \\ \uparrow & & \uparrow & & \uparrow & & \\ \operatorname{Hom}(M,E^2) & \longrightarrow & \operatorname{Hom}_R(P_0,E^2) & \longrightarrow & \operatorname{Hom}_R(P_1,E^2) & \longrightarrow & \operatorname{Hom}_R(P_2,E^2) \longrightarrow \cdots \\ \uparrow & & \uparrow & & \uparrow & & \\ \operatorname{Hom}(M,E^2) & \longrightarrow & \operatorname{Hom}_R(P_0,E^1) & \longrightarrow & \operatorname{Hom}_R(P_1,E^1) & \longrightarrow & \operatorname{Hom}_R(P_2,E^1) \longrightarrow \cdots \\ \uparrow & & \uparrow & & \uparrow & & \\ \operatorname{Hom}(M,E^2) & \longrightarrow & \operatorname{Hom}_R(P_0,E^1) & \longrightarrow & \operatorname{Hom}_R(P_1,E^1) & \longrightarrow & \operatorname{Hom}_R(P_2,E^1) \longrightarrow \cdots \end{array}$$

and

$$D = \begin{array}{ccccccc} \vdots & & \vdots & & \vdots & & \\ \uparrow & & \uparrow & & \uparrow & & \\ \operatorname{Hom}_R(P_0,E^2) & \longrightarrow & \operatorname{Hom}_R(P_1,E^2) & \longrightarrow & \operatorname{Hom}_R(P_2,E^2) & \longrightarrow & \cdots \\ \uparrow & & \uparrow & & \uparrow & & \\ \operatorname{Hom}_R(P_0,E^1) & \longrightarrow & \operatorname{Hom}_R(P_1,E^1) & \longrightarrow & \operatorname{Hom}_R(P_2,E^1) & \longrightarrow & \cdots \\ \uparrow & & \uparrow & & \uparrow & & \\ \operatorname{Hom}_R(P_0,N) & \longrightarrow & \operatorname{Hom}_R(P_1,N) & \longrightarrow & \operatorname{Hom}_R(P_2,N) & \longrightarrow & \cdots \,. \end{array}$$

We obtained C from $\mathrm{Hom}(P,E)$ by adding in a column induced by π, and D by adding in a row induced by ε. Now we notice that

$$\mathrm{cone}(\mathrm{Hom}_R(P,N) \longrightarrow \mathrm{Tot}^{\oplus}(\mathrm{Hom}(P,E))) = \mathrm{Tot}^{\oplus}(C),$$

and

$$\mathrm{cone}(\mathrm{Hom}_R(M,E) \longrightarrow \mathrm{Tot}^{\oplus}(\mathrm{Hom}(P,E))) = \mathrm{Tot}^{\oplus}(D).$$

The dual of Theorem 6.25 says that $\mathrm{Tot}^{\oplus}(C)$ and $\mathrm{Tot}^{\oplus}(D)$ are both exact, and thus $\mathrm{Hom}_R(P,N) \longrightarrow \mathrm{Tot}^{\oplus}(\mathrm{Hom}(P,E))$ and $\mathrm{Hom}_R(M,E) \longrightarrow \mathrm{Tot}^{\oplus}(\mathrm{Hom}(P,E))$ are both quasi-isomorphisms. We conclude that

$$R^n \mathrm{Hom}_R(M,-)(N) = \mathrm{H}^n(\mathrm{Hom}_R(P,N)) \cong \mathrm{H}^n(\mathrm{Hom}_R(M,E)) = R^n \mathrm{Hom}_R(-,N)(M). \quad \square$$

Definition 6.29. Let R be a ring and M and N be R-modules. The ith Tor module from M to N is

$$\mathrm{Tor}_i^R(M,N) := L_i(M \otimes_R -)(N) \cong L_i(- \otimes_R N)(M).$$

Notice in particular that the R-module $\mathrm{Tor}_i^R(M,N)$ is defined only up to isomorphism.

Definition 6.30. Let R be a ring and M and N be R-modules. The ith Ext module from M to N is

$$\mathrm{Ext}^i_R(M,N) := R^i \mathrm{Hom}_R(M,-)(N) \cong R^i \mathrm{Hom}_R(-,N)(M).$$

Notice in particular that the R-module $\mathrm{Ext}^i_R(M,N)$ is only defined up to isomorphism.

Theorem 6.6 immediately gives us long exact sequences for Ext and Tor.

Theorem 6.31. *Let R be a ring and M an R-module. Every short exact sequence of R-modules*

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

induces a long exact sequence

$$\cdots \longrightarrow \mathrm{Tor}^R_{n+1}(M,C) \longrightarrow \mathrm{Tor}^R_n(M,A) \longrightarrow \mathrm{Tor}^R_n(M,B) \longrightarrow \mathrm{Tor}^R_n(M,C) \longrightarrow \cdots$$
$$\cdots \longrightarrow \mathrm{Tor}^R_1(M,C) \longrightarrow A \otimes_R M \longrightarrow B \otimes_R M \longrightarrow C \otimes_R M \longrightarrow 0.$$

Theorem 6.32. *For every R-module M, every short exact sequence of R-modules*

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

induces a natural long exact sequence

$$0 \longrightarrow \mathrm{Hom}_R(M,A) \longrightarrow \mathrm{Hom}_R(M,B) \longrightarrow \mathrm{Hom}_R(M,C) \longrightarrow \mathrm{Ext}^1_R(M,A) \longrightarrow \cdots$$
$$\cdots \longrightarrow \mathrm{Ext}^n_R(M,B) \longrightarrow \mathrm{Ext}^n_R(M,C) \longrightarrow \mathrm{Ext}^{n+1}_R(M,A) \longrightarrow \cdots.$$

and

$$0 \longrightarrow \mathrm{Hom}_R(C,M) \longrightarrow \mathrm{Hom}_R(B,M) \longrightarrow \mathrm{Hom}_R(A,M) \longrightarrow \mathrm{Ext}^1_R(C,M) \longrightarrow \cdots$$
$$\cdots \longrightarrow \mathrm{Ext}^n_R(B,M) \longrightarrow \mathrm{Ext}^n_R(A,M) \longrightarrow \mathrm{Ext}^{n+1}_R(C,M) \longrightarrow \cdots.$$

Theorem 6.33. *Let M and N be R-modules. For all i, there are natural isomorphisms*

$$\operatorname{Tor}_i^R(M,N) \cong \operatorname{Tor}_i^R(N,M).$$

Proof. Let P be a projective resolution of M. By Theorem 6.27, $\operatorname{Tor}_i^R(M,N) = \operatorname{H}_i(P \otimes_R N)$ and $\operatorname{Tor}_i^R(N,M) = \operatorname{H}_i(N \otimes_R P)$. By Theorem 3.53, $M \otimes_R N$ and $N \otimes_R M$ are naturally isomorphic. In fact, $m \otimes n \mapsto n \otimes m$ determines an isomorphism. So consider the map

$$\begin{array}{ccc} P_n \otimes_R N \xrightarrow{f_n} N \otimes_R P_n & \qquad & N \otimes_R M \xrightarrow{g_n} P_n \otimes_R N \\ m \otimes n \longmapsto n \otimes m & & n \otimes m \longmapsto m \otimes n \end{array}$$

which again are isomorphisms for all n. Notice that these f_n assemble into a map of complexes $P \otimes_R N \xrightarrow{f} N \otimes_R P$, since

$$f_n(\partial(m \otimes n)) = f_n(\partial(m) \otimes n) = n \otimes \partial(m) = \partial(n \otimes m) = \partial f_{n+1}(m \otimes n).$$

Since all the f_n are isomorphisms, f is an isomorphism of complexes, and must then induce isomorphisms in homology. We conclude that

$$\operatorname{Tor}_i^R(M,N) = \operatorname{H}_i(P \otimes_R N) \cong \operatorname{H}_i(N \otimes_R P) = \operatorname{Tor}_i^R(N,M). \qquad \square$$

However, $\operatorname{Ext}^i_R(M,N)$ and $\operatorname{Ext}^i_R(N,M)$ can be dramatically different.

Example 6.34. Let k be a field and $R = k[x]$. The following is a minimal free resolution for $k = R/(x)$:

$$0 \longrightarrow R \xrightarrow{x} R \longrightarrow k \longrightarrow 0.$$

To compute $\operatorname{Ext}^i_R(k,R)$, we need only to apply $\operatorname{Hom}_R(-,R)$ to this resolution; one needs to be careful, though, as this is a contravariant functor. We obtain the following complex:

$$0 \longleftarrow \operatorname{Hom}_R(R,R) \xleftarrow{x^*} \operatorname{Hom}_R(R,R) \longleftarrow 0.$$

One can show that $\operatorname{Hom}_R(x,R)$ is multiplication by x on $\operatorname{Hom}_R(R,R)$; moreover, we have a natural isomorphism $\operatorname{Hom}_R(R,R) \cong R$, giving us

$$C = \qquad \begin{array}{ccccccc} 0 & \longleftarrow & R & \xleftarrow{x} & R & \longleftarrow & 0 \\ & & {\scriptstyle 1} & & {\scriptstyle 0} & & \end{array}$$

In particular,

$$\operatorname{Ext}^1(k,R) = \operatorname{H}^1(C) = R/(x).$$

In contrast,

$$\operatorname{Ext}^1(R,k) = 0$$

since R is free. Thus $\operatorname{Ext}^1(R,k) \ncong \operatorname{Ext}^1(k,R)$.

Exercise 6.35. Let M and N be R-modules with M finitely generated and let P be a prime ideal. Show that

$$\left(\operatorname{Ext}^i_R(M,N)\right)_P \cong \operatorname{Ext}^i_{R_P}(M_P, N_P).$$

Hint: reduce to the case when $i = 0$, which is Theorem 3.77.

There is an alternative description of Ext. It turns out that $\operatorname{Ext}^1_R(M,N)$ measures the **extensions** of M by N modulo split extensions. More precisely, an **extension** of M by N is a short exact sequence

$$0 \longrightarrow N \longrightarrow B \longrightarrow M \longrightarrow 0.$$

We can put an abelian group structure on the set of isomorphism classes of extensions of M by N, using an operation called the **Baer sum**, and one can show that the resulting abelian group is isomorphic to $\operatorname{Ext}^1_R(M,N)$. Via this description, the zero in $\operatorname{Ext}^1_R(M,N)$ corresponds to the split short exact sequence

$$0 \longrightarrow N \longrightarrow N \oplus M \longrightarrow M \longrightarrow 0.$$

The higher Ext modules can also be described in a similar fashion. First, we consider the set of n-fold extensions of N by M, meaning exact sequences of the form

$$0 \longrightarrow N \longrightarrow B_1 \longrightarrow B_2 \longrightarrow \cdots \longrightarrow B_n \longrightarrow M \longrightarrow 0$$

and the equivalence relation on this set given by the existence of a map of complexes

$$\begin{array}{ccccccccccc}
0 & \longrightarrow & N & \longrightarrow & B_1 & \longrightarrow & \cdots & \longrightarrow & B_n & \longrightarrow & M & \longrightarrow & 0 \\
 & & \| & & \downarrow & & & & \downarrow & & \| & & \\
0 & \longrightarrow & N & \longrightarrow & C_1 & \longrightarrow & \cdots & \longrightarrow & C_n & \longrightarrow & M & \longrightarrow & 0
\end{array}$$

where the vertical maps are not necessarily isomorphisms. We then define an operation on the set of equivalence classes of n-fold extensions of N by M that is also called the Baer sum, and one shows that the resulting abelian group is isomorphic to $\operatorname{Ext}^n_R(M,N)$.

Via this description, $\operatorname{Ext}^1_R(M,N) = 0$ if and only if every short exact sequence

$$0 \longrightarrow N \longrightarrow B \longrightarrow M \longrightarrow 0.$$

splits.

Finally, here are some nice facts about Ext and Tor we leave as exercises.

Exercise 6.36. If M and N are finitely generated R-modules and R is a noetherian ring, then $\operatorname{Ext}^i_R(M,N)$ and $\operatorname{Tor}^R_i(M,N)$ are both finitely generated R-modules for all i.

Exercise 6.37. Let R be a commutative ring and M and N be R-modules. Consider the R-module homomorphism $f\colon M \to M$ given by multiplication by a fixed element $r \in R$.

a) Show that the map $\operatorname{Tor}^R_i(f,M)\colon \operatorname{Tor}^R_i(M,N) \to \operatorname{Tor}^R_i(M,N)$ induced by f is multiplication by r on $\operatorname{Tor}^R_i(M,N)$.

b) Show that $\operatorname{Ext}^i(f,M)\colon \operatorname{Ext}^i_R(M,N) \to \operatorname{Ext}^i_R(M,N)$ is multiplication by r on $\operatorname{Ext}^i_R(M,N)$.

c) Show that the map $\operatorname{Ext}^i(M,f)\colon \operatorname{Ext}^i_R(N,M) \to \operatorname{Ext}^i_R(N,M)$ induced by f is multiplication by r on $\operatorname{Ext}^i_R(N,M)$.

Exercise 6.38. Let M be an R-module.

a) Show that M is flat if and only if $\operatorname{Tor}^R_1(M,N) = 0$ for every R-module N.

b) Show that M is projective if and only if $\operatorname{Ext}^1_R(M,N) = 0$ for every R-module N.

c) Show that M is injective if and only if $\operatorname{Ext}^1_R(N,M) = 0$ for every R-module N.

6.3 Computing Ext and Tor

Given R-modules M and N, we have two possible ways to compute $\operatorname{Tor}_i^R(M,N)$ from the definition.

Construction 6.39. Find a projective resolution

$$\cdots \longrightarrow P_2 \longrightarrow P_1 \longrightarrow P_0 \longrightarrow M \longrightarrow 0$$

of M. Applying $- \otimes_R N$ to the resolution (not counting M), we get a complex

$$\cdots \longrightarrow P_2 \otimes_R N \longrightarrow P_1 \otimes_R N \longrightarrow P_0 \otimes_R N \longrightarrow 0\,.$$

Its homology is $\operatorname{Tor}_*^R(M,N)$:

$$\operatorname{Tor}_i^R(M,N) = \operatorname{H}_i\left(\cdots \longrightarrow P_2 \otimes_R N \longrightarrow P_1 \otimes_R N \longrightarrow P_0 \otimes_R N \longrightarrow 0\right).$$

Alternatively, we can find a free resolution of N, say

$$\cdots \longrightarrow Q_2 \longrightarrow Q_1 \longrightarrow Q_0 \longrightarrow N \longrightarrow 0,$$

apply $M \otimes_R -$,

$$\cdots \longrightarrow M \otimes_R Q_2 \longrightarrow M \otimes_R Q_1 \longrightarrow M \otimes_R Q_0 \longrightarrow 0\,,$$

and compute the homology of the resulting complex:

$$\operatorname{Tor}_i^R(M,N) = \operatorname{H}_i\left(\cdots \longrightarrow M \otimes_R Q_2 \longrightarrow M \otimes_R Q_1 \longrightarrow M \otimes_R Q_0 \longrightarrow 0\right).$$

Similarly, we have two possible ways to compute $\operatorname{Ext}_R^i(M,N)$.

Construction 6.40. Find a projective resolution

$$\cdots \longrightarrow P_2 \longrightarrow P_1 \longrightarrow P_0 \longrightarrow M \longrightarrow 0$$

of M. Applying the contravariant functor $\operatorname{Hom}_R(-,N)$ to the resolution gives us a cocomplex rather than a complex:

$$0 \longrightarrow \operatorname{Hom}_R(P_0,N) \longrightarrow \operatorname{Hom}_R(P_1,N) \longrightarrow \operatorname{Hom}_R(P_2,N) \longrightarrow \cdots.$$

Its homology is $\operatorname{Ext}_R^*(M,N)$:

$$\operatorname{Ext}_R^i(M,N) = \operatorname{H}^i\left(0 \longrightarrow \operatorname{Hom}_R(P_0,N) \longrightarrow \operatorname{Hom}_R(P_1,N) \longrightarrow \operatorname{Hom}_R(P_2,N) \longrightarrow \cdots\right).$$

Alternatively, we can find an injective resolution of N, say

$$0 \longrightarrow N \longrightarrow E^0 \longrightarrow E^1 \longrightarrow E^2 \longrightarrow \cdots,$$

apply the covariant functor $\operatorname{Hom}_R(M,-)$, which yields the cocomplex

$$0 \longrightarrow \operatorname{Hom}_R(M,E^0) \longrightarrow \operatorname{Hom}_R(M,E^1) \longrightarrow \operatorname{Hom}_R(M,E^2) \longrightarrow \cdots,$$

and compute the cohomology of the resulting cocomplex:

$$\operatorname{Ext}_R^i(M,N) = \operatorname{H}^i\left(0 \longrightarrow \operatorname{Hom}_R(M,E^0) \longrightarrow \operatorname{Hom}_R(M,E^1) \longrightarrow \operatorname{Hom}_R(M,E^2) \longrightarrow \cdots\right).$$

It helps to keep a few simple ideas in mind:

- If P is a projective R-module, then
$$\mathrm{Tor}_i^R(M,P) = \mathrm{Tor}_i^R(P,M) = 0$$
and
$$\mathrm{Ext}^i_R(P,M) = 0$$
for all $i > 0$ and all R-modules M, since $0 \longrightarrow P \longrightarrow 0$ is a projective resolution for P. This is a special case of Theorem 6.3.

- If E is an injective R-module,
$$\mathrm{Ext}^i_R(M,E) = 0$$
for all $i > 0$ and all R-modules M.

- Free resolutions are often easier to compute explicitly, and the best path towards finding $\mathrm{Ext}^n_R(M,N)$.

- Relating one of our modules to other, easier modules via a short exact sequence can often simplify complicated computations.

Let's compute some examples.

Example 6.41. Let's compute $\mathrm{Ext}^i_{\mathbb{Z}}(\mathbb{Z}/(2),\mathbb{Z}/(3))$. Injective resolutions are not so easy to find, so we start from a projective resolution for $\mathbb{Z}/(2)$:

$$0 \longrightarrow \mathbb{Z} \xrightarrow{2} \mathbb{Z} \longrightarrow \mathbb{Z}/(2) \longrightarrow 0.$$

Notice that $\mathrm{pdim}_{\mathbb{Z}}(\mathbb{Z}/(2)) \neq 0$, since $\mathbb{Z}/(2)$ is not a projective $\mathbb{Z}$-module. We found a free resolution of length 1 for $\mathbb{Z}/(2)$, so it must be that $\mathrm{pdim}_{\mathbb{Z}}(\mathbb{Z}/(2)) = 1$. This immediately tells us that $\mathrm{Ext}^i_{\mathbb{Z}}(\mathbb{Z}/(2),\mathbb{Z}/(3)) = 0$ for all $i \leqslant 2$. Now we apply $\mathrm{Hom}_{\mathbb{Z}}(-,\mathbb{Z}/(3))$ to our free resolutions for $\mathbb{Z}/(2)$, and obtain

$$0 \longrightarrow \underset{0}{\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z},\mathbb{Z}/(3))} \xrightarrow{2^*} \underset{1}{\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z},\mathbb{Z}/(3))} \longrightarrow 0.$$

By Theorem 3.4, $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z},\mathbb{Z}/(3)) \cong \mathbb{Z}/(3)$, via the isomorphism $f \mapsto f(1)$. Since 2^* was the map $f \mapsto (2\cdot -)\circ f = 2f(-)$, we can simplify our complex to

$$0 \longrightarrow \mathbb{Z}/(3) \xrightarrow{2} \mathbb{Z}/(3) \longrightarrow 0.$$

Notice that multiplication by 2 is an isomorphism on $\mathbb{Z}/(3)$, so the complex above is exact, and $\mathrm{Ext}^i_{\mathbb{Z}}(\mathbb{Z}/(2),\mathbb{Z}/(3)) = 0$ for all i.

Example 6.42. Given an integer $n > 1$,

$$0 \longrightarrow \mathbb{Z} \xrightarrow{n} \mathbb{Z} \xrightarrow{\pi} \mathbb{Z}/(n) \longrightarrow 0$$

with π the canonical projection is a free resolution for $\mathbb{Z}/(n)$ over $\mathbb{Z}$. Notice that since $\mathbb{Z}/(n)$ is not a free $\mathbb{Z}$-module, there is no shorter free resolution for $\mathbb{Z}/\mathfrak{n}$. Now we can use this resolution to compute $\mathrm{Tor}_i^{\mathbb{Z}}(\mathbb{Z}/(n), M)$ and $\mathrm{Ext}^i_{\mathbb{Z}}(\mathbb{Z}/(n), M)$ for any $\mathbb{Z}$-module M. For Tor,

$$\mathrm{Tor}_i^{\mathbb{Z}}(\mathbb{Z}/(n), M) = \mathrm{H}_i(\, 0 \longrightarrow \mathbb{Z} \otimes_{\mathbb{Z}} M \xrightarrow{n\otimes 1} \mathbb{Z} \otimes_{\mathbb{Z}} M \longrightarrow 0\,).$$

By Theorem 3.55 $\mathbb{Z} \otimes_{\mathbb{Z}} M \cong M$, via the map $k \otimes m \mapsto km$, and the map $n \otimes 1_M$ corresponds to multiplication by n on M. Therefore,

$$\mathrm{Tor}_i^{\mathbb{Z}}(\mathbb{Z}/(n), M) = \mathrm{H}_i(\, 0 \longrightarrow M \xrightarrow{n} M \longrightarrow 0\,),$$

so

$$\mathrm{Tor}_i^{\mathbb{Z}}(\mathbb{Z}/(n), M) = \begin{cases} M/nM & \text{for } i = 0 \\ (0 :_M n) & \text{for } i = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $\mathrm{Tor}_0^{\mathbb{Z}}(\mathbb{Z}/(n), M) = M/nM = \mathbb{Z}/n\mathbb{Z} \otimes_{\mathbb{Z}} M$, as we already knew from Theorem 6.4.

Similarly, we can compute all the Ext modules from $\mathbb{Z}/(n)$:

$$\mathrm{Ext}^i_{\mathbb{Z}}(\mathbb{Z}/(n), M) = \mathrm{H}^i(\, 0 \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}, M) \xrightarrow{n^*} \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}, M) \longrightarrow 0\,).$$

By Theorem 3.4, $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}, M) \cong M$, via the map $f \mapsto f(1)$, and $n^* = \mathrm{Hom}_{\mathbb{Z}}(n, M)$ corresponds to multiplication by n on M. So

$$\mathrm{Ext}^i_{\mathbb{Z}}(\mathbb{Z}/(n), M) = \mathrm{H}^i(\, 0 \longrightarrow M \xrightarrow{n} M \longrightarrow 0\,).$$

We conclude that

$$\mathrm{Ext}^i_{\mathbb{Z}}(\mathbb{Z}/(n), M) = \begin{cases} M/nM & \text{for } i = 1 \\ (0 :_M n) & \text{for } i = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $\mathrm{Ext}^1_{\mathbb{Z}}(\mathbb{Z}/(n), M) = (0 :_M n) = \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/(n), M)$, as we already knew from Theorem 6.4.

Alternatively, we can compute $\mathrm{Ext}^i_{\mathbb{Z}}(\mathbb{Z}/(n), M)$ and $\mathrm{Tor}_i^{\mathbb{Z}}(\mathbb{Z}/(n), M)$ by looking at some long exact sequences. The long exact sequence for Tor induced by the short exact sequence

$$0 \longrightarrow \mathbb{Z} \xrightarrow{n} \mathbb{Z} \longrightarrow \mathbb{Z}/(n) \longrightarrow 0$$

is

$$\cdots \longrightarrow \mathrm{Tor}_{n+1}^{\mathbb{Z}}(\mathbb{Z}/(n), M) \longrightarrow \mathrm{Tor}_n^{\mathbb{Z}}(\mathbb{Z}, M) \longrightarrow \mathrm{Tor}_n^{\mathbb{Z}}(\mathbb{Z}, M) \longrightarrow \mathrm{Tor}_n^{\mathbb{Z}}(\mathbb{Z}/(n), M) \longrightarrow \cdots$$
$$\cdots \longrightarrow \mathrm{Tor}_1^{\mathbb{Z}}(\mathbb{Z}/(n), M) \longrightarrow \mathbb{Z} \otimes_{\mathbb{Z}} M \longrightarrow \mathbb{Z} \otimes_{\mathbb{Z}} M \longrightarrow \mathbb{Z}/(n) \otimes_{\mathbb{Z}} M \longrightarrow 0.$$

Since $\mathbb{Z}$ is a projective $\mathbb{Z}$-module and thus flat, $\operatorname{Tor}_i^{\mathbb{Z}}(\mathbb{Z}, M) = 0$ for all $i > 0$. As a consequence, the long exact sequence above forces $\operatorname{Tor}_2^{\mathbb{Z}}(\mathbb{Z}/(n), M) = 0$. So our long exact sequence really gets reduced to

$$0 \longrightarrow \operatorname{Tor}_1^{\mathbb{Z}}(\mathbb{Z}/(n), M) \longrightarrow \mathbb{Z} \otimes_{\mathbb{Z}} M \longrightarrow \mathbb{Z} \otimes_{\mathbb{Z}} M \longrightarrow \mathbb{Z}/(n) \otimes_{\mathbb{Z}} M \longrightarrow 0.$$

Now $\mathbb{Z} \otimes_{\mathbb{Z}} M \cong M$ via $k \otimes m \mapsto km$, and this isomorphism turns $n \otimes 1_M$ into multiplication by n on M, same as above. So $\operatorname{Tor}_1^{\mathbb{Z}}(\mathbb{Z}/(n), M)$ is the kernel of multiplication by n on M, or $(0 :_M n)$.

If we want to compute $\operatorname{Ext}^i_{\mathbb{Z}}(\mathbb{Z}/(n), M)$, we should now look at the long exact sequence

$$0 \longrightarrow \operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}/(n), M) \longrightarrow \operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}, M) \xrightarrow{n^*} \operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}, M) \longrightarrow \operatorname{Ext}^1_{\mathbb{Z}}(\mathbb{Z}/(n), M) \longrightarrow \cdots$$
$$\cdots \longrightarrow \operatorname{Ext}^n_{\mathbb{Z}}(\mathbb{Z}, M) \longrightarrow \operatorname{Ext}^n_{\mathbb{Z}}(\mathbb{Z}, M) \longrightarrow \operatorname{Ext}^{n+1}_{\mathbb{Z}}(\mathbb{Z}/(n), M) \longrightarrow \cdots.$$

Again, $\mathbb{Z}$ is a free $\mathbb{Z}$-module, so $\operatorname{Ext}^i_{\mathbb{Z}}(\mathbb{Z}, M) = 0$ for all $i > 0$. Then $\operatorname{Ext}^i_{\mathbb{Z}}(\mathbb{Z}/(n), M) = 0$ for all $i > 1$, and our long exact sequence is actually just

$$0 \longrightarrow \operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}/(n), M) \longrightarrow \operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}, M) \xrightarrow{n^*} \operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}, M) \longrightarrow \operatorname{Ext}^1_{\mathbb{Z}}(\mathbb{Z}/(n), M) \longrightarrow 0.$$

So $\operatorname{Ext}^1_{\mathbb{Z}}(\mathbb{Z}/(n), M)$ is the cokernel of n^*. As before, notice that $\operatorname{Hom}_{\mathbb{Z}}(\mathbb{Z}, M) \cong M$ via the map $f \mapsto f(1)$, and n^* corresponds to multiplication by n on M. We conclude that $\operatorname{Ext}^1_{\mathbb{Z}}(\mathbb{Z}/(n), M) \cong M/nM$.

Exercise 6.43. Let k be a field and $R = k[x]/(x^3)$.

a) Compute $\operatorname{Tor}_i^R(k, k)$ for all i.

b) Show that $\operatorname{Ext}^i_R(k, k) \neq 0$ for all i.

Exercise 6.44. Let k be a field, $R = k[x, y]$, and $\mathfrak{m} = (x, y)$.

a) Show that

$$0 \longrightarrow R \xrightarrow{\begin{pmatrix} y \\ -x \end{pmatrix}} R^2 \xrightarrow{\begin{pmatrix} x & y \end{pmatrix}} R \longrightarrow 0$$

is a free resolution for $k = R/\mathfrak{m}$.

b) Compute $\operatorname{Tor}_i^R(k, k)$ for all i.

c) Show that

$$\operatorname{Tor}_1(\mathfrak{m}, k) \cong \operatorname{Tor}_2(k, k).$$

6.4 Other derived functors

Here are some other examples of derived functors you may encounter.

Group homology and group cohomology

Definition 6.45. category of G-modules Let G be a group. A **(left) G-module** is an abelian group A with an action of G by additive maps on the left, meaning that

$$g(a+b) = ga + gb$$

for all $a, b \in A$ and all $g \in G$, where we write ga for the action of $g \in G$ on $a \in A$. Given two G-modules A and B, a morphism of G-modules $f : A \to B$ is a group homomorphism that is also **G-equivariant**, meaning $f(ga) = gf(a)$ for all $g \in G$ and $a \in A$.

The **category of G-modules**, which we write as **G-mod**, has objects all G-modules and arrows all G-module morphisms. We write $\mathrm{Hom}_G(A, B)$ instead of $\mathrm{Hom}_{G\text{-}\mathbf{mod}}(A, B)$.

This category **G-mod** can be identified with the category of $\mathbb{Z}[G]$-modules, of modules over the (noncommutative) ring $\mathbb{Z}G$, the group ring of G. It can also be identified with the functor category $\mathbf{Ab}^G$ of functors from the category G to the category **Ab** of abelian groups. As a reminder, G gives a category with one object G and arrows the elements of G.

Definition 6.46. The **invariant subgroup** A^G of a G-module A is

$$A^G := \{a \in A \mid ga = a \text{ for all } g \in G\}.$$

The **coinvariant subgroup** A_G of a G-module A is

$$A_G := A/G\text{-submodule generated by } \{ga - a \in A \mid g \in G, a \in A\}.$$

Exercise 6.47. Given any G-module A, $A_G \cong \mathbb{Z} \otimes_{\mathbb{Z}G} A$ and $A^G \cong \mathrm{Hom}_G(\mathbb{Z}, A)$, where $\mathbb{Z}$ denotes the trivial G-module. In fact, there are natural isomorphisms $(-)_G \cong \mathbb{Z} \otimes_{\mathbb{Z}G} -$ and $(-)^G \cong \mathrm{Hom}_G(\mathbb{Z}, -)$.

Thus taking coinvariants is right exact, and taking invariants is left exact.

Definition 6.48. Let G be a group and A a G-module. Group homology is the derived functor $\mathrm{H}_i(G; -) := L_i(-_G)$; the **homology groups of G with coefficients in A** are

$$\mathrm{H}_i(G; A) := L_i(-_G)(A).$$

Group cohomology is the derived functor $\mathrm{H}^i(G; -) := R^i(-^G)$; the **cohomology groups of G with coefficients in A** are the G-modules:

$$\mathrm{H}^i(G; A) := R^i(-^G)(A).$$

By Theorem 6.47,

$$\mathrm{H}_i(G; A) \cong \mathrm{Tor}_i^{\mathbb{Z}[G]}(\mathbb{Z}, A) \text{ and } \mathrm{H}^i(G; A) \cong \mathrm{Ext}^i_{\mathbb{Z}[G]}(\mathbb{Z}, A).$$

Thus to compute group (co)homology we need a projective resolution for the trivial $\mathbb{Z}[G]$-module $\mathbb{Z}$. Note also that by Theorem 6.4, $\mathrm{H}_0(G; A) = A_G$ and $\mathrm{H}^0(G; A) = A^G$.

Group (co)homology is a rich subject. For a detailed treatment of group (co)homology, see Weibel's *Homological Algebra* [Wei94].

Local Cohomology

Let I be an ideal in a ring R. The **I-torsion** functor $\Gamma_I \colon R\text{-}\mathbf{Mod} \longrightarrow R\text{-}\mathbf{Mod}$ is defined by

$$\Gamma_I(M) := \{m \in M \mid I^n m = 0 \text{ for some } n\}$$

which acts on maps by restriction.

Exercise 6.49. The I-torsion functor is a left exact covariant additive functor.

The I-torsion functor gives rise to local cohomology, the right derived functors H^i_I of Γ_I. The ith local cohomology of M with support on I is then given by

$$\mathrm{H}^i_I(M) := R^i\Gamma_I(M).$$

Local cohomology was introduced by Grothendieck in a series of seminars at Harvard in 1961, which are now of course very famous. Grothendieck himself never published any notes on the subject, but Robin Hartshorne's notes of those lectures have been published.

Local cohomology is a rich subject, and we could easily spend an entire semester on it. For a modern treatment of the local cohomology and its connections, the book *24 hours of local cohomology* [ILL+07] and the very nice notes by Craig Huneke, Mel Hochster, and Jack Jeffries are all excellent resources.

It turns out that local cohomology modules can be defined in a few different ways, which are in no way obviously equivalent, and those different points of view are quite helpful. For example, we can define local cohomology via the Čech complex.

Definition 6.50 (Čech complex)**.** Let M be an R-module and $x \in R$. The **Čech complex** of x on R is given by

$$\check{C}^\bullet(x) := \left(0 \longrightarrow R \underset{0}{\longrightarrow} R_x \underset{1}{\longrightarrow} 0 \right)$$

The **Čech complex** of $f_1, \ldots, f_t \in R$ on M is given by

$$\check{C}^\bullet(f_1^n, \ldots, f_t^n; M) := \check{C}^\bullet(f_1) \otimes \cdots \otimes \check{C}^\bullet(f_t) \otimes M.$$

Example 6.51. Let's compute the Čech complex on f and g and an R-module M.

$$\check{C}^\bullet(f, g; M) \quad = \quad \begin{array}{ccccc} 0 & & 0 & & \\ \uparrow & & \uparrow & & \\ 0 & \longrightarrow & M_g & \xrightarrow[-1]{} & M_{fg} \longrightarrow 0 \\ & & \uparrow{\scriptstyle 1} & & \uparrow{\scriptstyle 1} \\ & & M & \xrightarrow[1]{} & M_f \longrightarrow 0 \\ & & \uparrow & & \uparrow \\ & & 0 & & 0 \end{array} \quad = \quad 0 \longrightarrow M \xrightarrow{\binom{1}{1}} M_f \oplus M_g \xrightarrow{(1\ -1)} M_{fg} \longrightarrow 0$$

2

1

0

Exercise 6.52.

a) $\check{C}^\bullet(f_1,\dots,f_t;M) \cong \bigoplus_{\{j_1,\dots,j_i\}\subseteq[t]} M_{f_{j_1}\cdots f_{j_i}}$

b) The maps between components corresponding to subsets I, J are zero if $I \not\subseteq J$, and ± 1 if $J = I \cup \{k\}$.

It turns out that the cohomology of the Čech complex gives us local cohomology. For an ideal $I = (f_1,\dots,f_n)$,

$$\begin{aligned}\mathrm{H}^i_I(M) &= \mathrm{H}^i(\check{C}^\bullet(f_1,\dots,f_n;M)) \\ &= \mathrm{H}^i\left(0 \to M \to \cdots \to \bigoplus_i M_{f_i} \to \cdots \bigoplus_{i=1}^n M_{f_1\cdots\widehat{f_i}\cdots f_n} \to M_{f_1\cdots f_n} \to 0\right)\end{aligned}$$

so elements in the ith local cohomology can be realized as equivalence classes of fractions.

Local cohomology modules also arise as a direct limit of Ext modules:

$$\varinjlim_n \mathrm{Ext}^i_R(R/I^n, M)$$

The equivalence between all these different definitions is a fundamental result in the theory of local cohomology.

Local cohomology modules play a crucial, ubiquitous role in commutative algebra. They measure many important invariants, such as dimension and depth, and are extremely useful tools for studying all sorts of topics; for example, they can be used to detect if a ring is Gorenstein (if it has finite injective dimension as a module over itself) or Cohen-Macaulay (a nice class of rings that is both very large but also very well behaved). However, local cohomology modules are typically not finitely generated. One reason for this is that injective modules are also often not finitely generated. Local cohomology is also a major reason why commutative algebraists are interested in studying injective modules.

In fact, local cohomology is almost *never* finitely generated. Here's a very simple example.

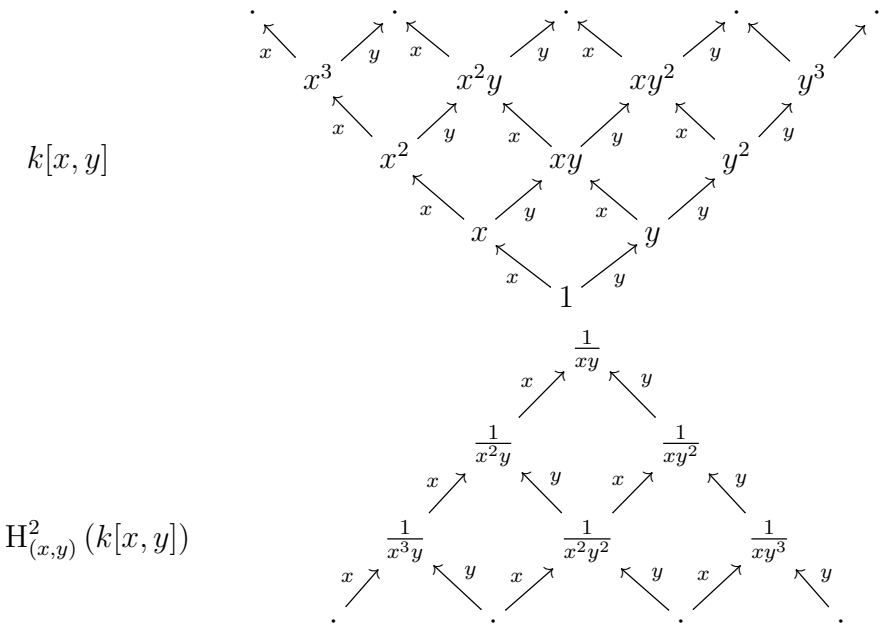
Example 6.53. Let $R = k[x_1,\dots,x_n]$, k be a field, and $\mathfrak{m} = (x_1,\dots,x_n)$. Then $\mathrm{H}^n_{\mathfrak{m}}(R)$ has the k-vector space structure

$$\bigoplus_{\text{all } a_i>0} k \cdot \frac{1}{x_1^{a_1}\cdots x_n^{a_n}},$$

with R-module structure given by

$$x_1^{b_1}\cdots x_n^{b_n} \cdot \frac{z}{x_1^{a_1}\cdots x_n^{a_n}} = \begin{cases} \frac{z}{x_1^{a_1-b_1}\cdots x_n^{a_n-b_n}} & \text{if all } b_i < a_i \\ 0 & \text{otherwise.}\end{cases}$$

This is not a finitely generated module! Note also that every finitely generated submodule only has terms with bounded negative degree. But this is still a very nice module: it looks like R upside down.



Despite being infinitely generated, local cohomology modules enjoy many finiteness properties we have gotten used to expecting from finitely generated modules. For example, over a local ring $(R, \mathfrak{m})$, the local cohomology modules $\mathrm{H}^i_{\mathfrak{m}}(M)$ of a finitely generated module M are Artinian — but not Noetherian!

Huneke raised the question of whether local cohomology modules of noetherian rings always have finitely many associated primes, a problem which has been a very active research are in commutative algebra in the last few decades. While the answer to Huneke's question is no — as famous examples by Katzmann, Singh, and Singh and Swanson show — the local cohomology modules of finitely generated R-modules over a regular ring do have finitely many associated primes.

One very important invariant we can study with local cohomology is the arithmetic rank.

Definition 6.54. Let I be an ideal in a Noetherian ring R. The **arithmetic rank** of I is defined by

$$\operatorname{ara}(I) := \min\{s \mid \text{there exist some } x_1, \ldots, x_s \text{ such that } \sqrt{(x_1, \ldots, x_s)} = \sqrt{I}\}.$$

Given a variety $X = V(I) \subseteq \mathbb{A}^n_k$, the arithmetic rank of its defining ideal $I(X)$ is the minimum number of equations needed to define X. It turns out that this number is difficult to study, and it is best understood via local cohomology, a thought best described by Lyubeznik:

> Part of what makes the problem about the number of defining equations so interesting is that it can be very easily stated, yet a solution, in those rare cases when it is known, usually is highly nontrivial and involves a fascinating interplay of Algebra and Geometry.
>
> (Lyubeznik, in [Lyu92])

The connection to local cohomology begins with the following two elementary facts about local cohomology:

- If $\sqrt{I} = \sqrt{J}$, then $\mathrm{H}^i_I(-) = \mathrm{H}^i_J(-)$.
- Given any ideal I, $\operatorname{ara}(I) \geqslant \min\{i \mid \mathrm{H}^i_I(M) \neq 0 \text{ for some } R\text{-module } M\}$.

So computing local cohomology modules, or deciding when they vanish, can help us find bounds on the arithmetic rank of a variety.

We close this chapter with yet another example of a derived functor of an interesting functor.

Exercise 6.55. Let R be a domain and Q be its fraction field. Let T denote the torsion functor.

a) Show that $T(M) = \operatorname{Tor}^R_1(M, Q/R)$.

b) Show that for every short exact sequence

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

of R-modules gives rise to an exact sequence

$$0 \longrightarrow T(A) \longrightarrow T(B) \longrightarrow T(C) \longrightarrow (Q/R) \otimes_R A \longrightarrow (Q/R) \otimes_R B \longrightarrow (Q/R) \otimes_R C \longrightarrow 0.$$

c) Show that the right derived functors of T are $R^1T = (Q/R) \otimes_R -$ and $R^iT = 0$ for all $i \geqslant 2$.

Chapter 7

Abelian categories

An abelian category is a category that has just enough extra structure to behave like R-**Mod**: we have complexes and exact sequences, homology, the Snake Lemma, the long exact sequence in homology, and many other nice features. On the one hand, every abelian category embeds nicely in some R-**Mod**, so it is in some ways sufficient to study R-**Mod**. In other ways, the general nonsense definitions in an abelian category can sometimes give us a uniform, simple way to prove many results about R-**Mod** (and Ch(R-**Mod**), and other related categories) all at once.

7.1 What's an abelian category?

Definition 7.1. A category $\mathscr{A}$ is a **preadditive category** if:

- For all objects x and y in $\mathcal{A}$, $\mathrm{Hom}_{\mathcal{A}}(x, y)$ is an abelian group.
- For all objects x, y, and z in $\mathcal{A}$, the composition

$$\mathrm{Hom}_{\mathcal{A}}(x, y) \times \mathrm{Hom}_{\mathcal{A}}(z, x) \xrightarrow{\circ} \mathrm{Hom}_{\mathcal{A}}(z, y)$$

  is bilinear, meaning

$$g \circ (f_1 + f_2) = g \circ f_1 + g \circ f_2 \quad \text{and} \quad (g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f.$$

In the literature, preadditive categories are sometimes called **Ab-enriched categories**.

Example 7.2. Our favorite category R-**Mod** is a preadditive category; so is Ch(R).

We can talk about additive functors between any two preadditive categories.

Definition 7.3. Let $\mathcal{A}$ and $\mathscr{B}$ be preadditive categories. An **additive functor** $\mathcal{A} \to \mathscr{B}$ is a functor such that the map

$$\begin{aligned} \mathrm{Hom}_{\mathcal{A}}(x, y) &\longrightarrow \mathrm{Hom}_{\mathscr{B}}(F(x), F(y)) \\ f &\longmapsto F(f) \end{aligned}$$

is a homomorphism of abelian groups.

Recall the notions of initial, terminal, and zero objects, which we discussed in Chapter 1.

Definition 7.4. Let $\mathscr{C}$ be a category with a zero object 0. Given two objects x and y in $\mathscr{C}$, the **zero arrow** from x to y is the unique arrow $x \longrightarrow y$ that factors through 0, meaning the arrow given by composition of the unique arrows $x \longrightarrow 0 \longrightarrow y$. We will often denote both the zero object and the zero arrow by 0, whenever it does not lead to confusion.

Remark 7.5. If a category $\mathcal{A}$ has a zero object, then $\mathrm{Hom}_{\mathcal{A}}(x, y)$ is always nonempty, since it contains at least the 0 arrow.

Remark 7.6. Composing the zero arrow with any other arrow always yields the zero arrow.

Remark 7.7. In any preadditive category $\mathcal{A}$ with a zero object 0, the 0 arrow $x \longrightarrow y$ coincides with the 0 of the abelian group $\mathrm{Hom}_{\mathcal{A}}(x, y)$.

Remark 7.8. We can characterize the zero object 0 by the property that the zero arrow and the identity arrows on 0 coincide. To see this, notice that if $1_x = x \xrightarrow{0} x$, then given any arrow $x \xrightarrow{f} y$, we must have

$$f = f \circ 1_x = f \circ 0 = 0,$$

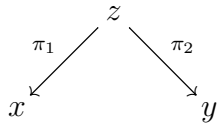
and similarly any arrow $y \xrightarrow{f} x$ must be 0. Then x is terminal and initial, and it must be the zero object.

Definition 7.9. An **additive category** is a preadditive category $\mathcal{A}$ such that:

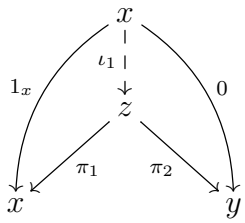
- $\mathcal{A}$ has a zero object.
- $\mathcal{A}$ has all finite products: given any two objects x and y in $\mathcal{A}$, there exists a product of x and y in $\mathcal{A}$.

Lemma 7.10. *In an additive category, finite coproducts exist and they agree with products.*

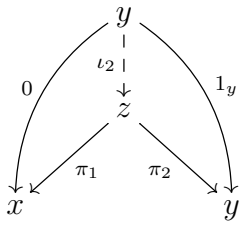
Proof. Let x and y be objects in our additive category, and consider their product, which exists by assumption:



The universal property of the product give arrows ι_1 and ι_2 such that

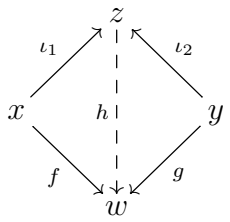


and



commute.

We claim that z together with ι_1 and ι_2 form a coproduct for x and y. Given an object w and arrows $x \xrightarrow{f} w$ and $y \xrightarrow{g} w$, we need to show that there exists a unique arrow h such that



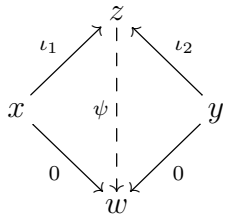
commutes.

To see such an h exists, consider $h := f\pi_1 + g\pi_2$. Then

$$h\iota_1 = f\underbrace{\pi_1\iota_1}_{1_x} + g\underbrace{\pi_2\iota_1}_{0} = f \quad \text{and} \quad h\iota_2 = f\underbrace{\pi_1\iota_2}_{0} + g\underbrace{\pi_2\iota_2}_{1_y} = f,$$

so indeed our proposed h does the job.

To show the uniqueness of such an h, we will use the fact that z together with π_1 and π_2 is a product for x and y. So suppose that h' is another arrow such that $h'\iota_1 = f$ and $h'\iota_2 = g$. Then $h - h'$ satisfies $(h - h')\iota_1 = f - f = 0$ and $(h - h')\iota_2 = g - g = 0$, so it's sufficient to show that the 0 arrow is the unique arrow ψ such that



commutes. First, we claim that $\iota_1\pi_1 + \iota_2\pi_2$ is the identity arrow on z. And indeed, this map satisfies

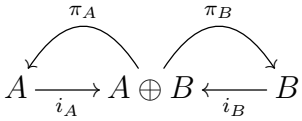
$$\pi_1(\iota_1\pi_1 + \iota_2\pi_2) = \underbrace{\pi_1\iota_1}_{1_x}\pi_1 + \underbrace{\pi_2\iota_1}_{0}\pi_2 = \pi_1 \quad \text{and} \quad \pi_2(\iota_1\pi_1 + \iota_2\pi_2) = \underbrace{\pi_2\iota_1}_{0}\pi_1 + \underbrace{\pi_2\iota_1}_{1_y}\pi_2 = \pi_2,$$

and so does the identity arrow 1_z, so the universal property of the product guarantees that $\iota_1\pi_1 + \iota_2\pi_2 = 1_z$. Now if $\psi\iota_1 = 0$ and $\psi\iota_2 = 0$, then

$$\psi = \psi 1_z = \psi(\iota_1\pi_1 + \iota_2\pi_2) = \psi\iota_1\pi + \psi\iota_2\pi_2 = 0 + 0 = 0. \qquad \square$$

Notation 7.11. In an additive category $\mathcal{A}$, given objects A and B, the notation $A \oplus B$ denotes the product $\equiv$coproduct of A and B.

Remark 7.12. If $\mathcal{A}$ is an additive category, the object $A \oplus B$ is characterized by the existence of arrows



such that

$$\pi_A i_A = \mathrm{id}_A, \quad \pi_B i_B = \mathrm{id}_B, \quad \text{and} \quad i_A\pi_A + i_B\pi_B = \mathrm{id}_{A\oplus B}\,.$$

Lemma 7.13. *Let $F : \mathcal{A} \longrightarrow \mathcal{B}$ an additive functor between additive categories, and let* 0 *denote the zero object on both categories.*

a) We have $F(0) = 0$. For any two objects x and y, $F(x \xrightarrow{0} y) = F(x) \xrightarrow{0} F(y)$.

b) F preserves finite products and coproducts.

Proof. We show the statement assuming F is covariant, and note that the argument in the contravariant case is essentially the same.

a) Since $F_{xy} : \mathrm{Hom}_{\mathcal{A}}(x, y) \to \mathrm{Hom}_{\mathcal{B}}(F(x), F(y))$ is a group homomorphism and that the zero elements in the abelian groups $\mathrm{Hom}_{\mathcal{A}}(x, y)$ and $\mathrm{Hom}_{\mathcal{B}}(F(x), F(y))$ are the corresponding zero arrows, then

$$F(x \xrightarrow{0} y) = F(x) \xrightarrow{0} F(y).$$

Now the zero arrow and the identity arrows of the zero object coincide, and so do their images by F. On the one hand, $F(1_0) = 1_{F(0)}$. On the other hand, by what we have shown at the arrow level we have

$$F(1_0) = F(0) \xrightarrow{0} F(0).$$

Then the identity and the zero arrows on $F(0)$ coincide, so by Theorem 7.8 we must have $F(0) = 0$.

b) Fix objects A and B and the canonical arrows

$$A \xrightarrow{i_A} A \oplus B, \quad B \xrightarrow{i_B} A \oplus B, \quad A \oplus B \xrightarrow{\pi_A} A, \quad \text{and} \quad A \oplus B \xrightarrow{\pi_A} B.$$

Any functor preserves identity arrows, so any additive functor F must satisfy

$$F(\pi_A)F(i_A) = F(\pi_A i_A) = \mathrm{id}_{F(A)} \qquad F(\pi_B)F(i_B) = F(\pi_B i_B) = \mathrm{id}_{F(B)}$$

and

$$F(i_A)F(\pi_A) + F(i_B)F(\pi_B) = \mathrm{id}_{F(A\oplus B)},$$

which satisfy

$$\pi_A i_A = \mathrm{id}_A, \quad \pi_B i_B = \mathrm{id}_B, \quad \text{and} \quad i_A\pi_A + i_B\pi_B = \mathrm{id}_{A\oplus B}\,.$$

By Theorem 7.12, this implies that $F(A \oplus B)$ is the product $\equiv$ coproduct of $F(A)$ and $F(B)$. □

Exercise 7.14. Let $\mathcal{A}$ be an additive category.

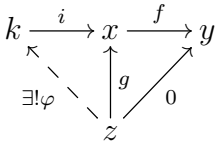
a) Show that an arrow f is a mono if and only if $fg = 0$ implies $g = 0$.

b) Show that an arrow f is an epi if and only if $gf = 0$ implies $g = 0$.

We can now define kernels and cokernels.

Definition 7.15. Let $\mathcal{A}$ be an additive category and f an arrow $x \longrightarrow y$. The **kernel** of f is an arrow $k \xrightarrow{i} x$ satisfying the following properties:

- $k \xrightarrow{i} x \xrightarrow{f} y$ is 0.
- Given any $g \in \mathrm{Hom}_{\mathcal{A}}(z, x)$ such that $z \xrightarrow{g} x \xrightarrow{f} y$ is the zero arrow, there exists a unique arrow φ such that $i\varphi = g$, meaning that

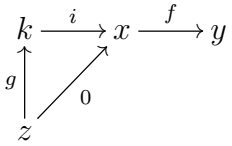


commutes. We denote the kernel of f by $\ker f$.

Remark 7.16. We claim that a kernel, if it exists, is always a mono. Indeed, suppose that

$$z \underset{g_2}{\overset{g_1}{\rightrightarrows}} k \xrightarrow{i} x \xrightarrow{f} y$$

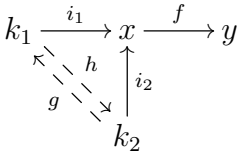
are such that $ig_1 = ig_2$. Then $i(g_1 - g_2) = 0$, so it's sufficient to show that $ig = 0$ implies $g = 0$. But then



commutes, and $f \circ 0 = 0$, so 0 factors uniquely through the kernel. But both g and $z \xrightarrow{0} k$ are such factorizations, so $g = 0$.

We are used to thinking about the kernel of a map f as an object; but in this general context, the kernel is really an arrow, or more precisely, an object (the source of the kernel) and an arrow from that object to the source of f. We sometimes refer to the kernel as the pair (object, arrow). Also, we might use the notation $\ker f \longrightarrow x$ for the kernel of $f : x \to y$. We might also abuse notation and refer to the object that is the source of $\ker f$ as the kernel of f, motivated by the familiar case of R-**Mod**. Nevertheless, the kernel of f is technically an arrow, not an object. A good reason for identifying the arrow $\ker f$ with its source object is the following rewriting of the definition:

Remark 7.17. If $k_1 \xrightarrow{i_1} x$ and $k_2 \xrightarrow{i_2} x$ are both kernels of f, then there exist unique arrows g and h such that

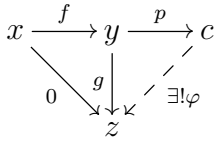


commutes. By Theorem 7.16, kernels are always monic. But then $i_1gh = i_2h = i_1$, and since i_1 is a mono, we must have $gh = 1$. Similarly, $hg = 1$, and g and h are isomorphisms.

This shows that if $k \xrightarrow{i} x$ is the kernel of $f \in \mathrm{Hom}_{\mathcal{A}}(x, y)$, the object k is, up to isomorphism, the unique object that satisfies the following universal property: for every object z and every arrow $g\colon z \to x$ such that $fg = 0$, there exists a unique arrow $h\colon z \to k$ such that $ih = g$.

Definition 7.18. Let $\mathcal{A}$ be an additive category and $f \in \mathrm{Hom}_{\mathcal{A}}(x, y)$. The **cokernel** of f is an arrow $y \xrightarrow{p} c$, denoted $\operatorname{coker} f$, satisfying the following properties:

- $x \xrightarrow{f} y \xrightarrow{p} c$ is 0.
- Given any $g \in \mathrm{Hom}_{\mathcal{A}}(y, z)$ such that $x \xrightarrow{f} y \xrightarrow{g} z$ is 0, there exists a unique arrow φ such that $\varphi p = g$, meaning that



  commutes.

We will sometimes use the notation $y \to \operatorname{coker} f$ for the cokernel of $x \xrightarrow{f} y$, although once again the cokernel of f is an arrow rather than an object.

Example 7.19.

a) The kernels and cokernels in R-**Mod** are what we think they are: the inclusion of the usual kernel, and the projection onto the usual cokernel.

b) It's not always true that all arrows have kernels or cokernels. For example, the category of finitely generated R-modules over some nonnoetherian ring R is additive, but it does not have all kernels. If I is some infinitely generated ideal in R, the kernel of the canonical projection $R \longrightarrow R/I$ does not exist in our category. In fact, this is an epi but not a cokernel: it should be the cokernel of the inclusion map $I \longrightarrow R$, but that is not an arrow in our category.

While not all epis are cokernels and not all monos are kernels, the converse is true. Just like we saw for kernels, cokernels, if they exist, are always epi, and they are unique in the sense we described in Theorem 7.17.

Exercise 7.20. Let $\mathcal{A}$ be an additive category.

a) Show if $y \xrightarrow{\pi} c$ is the cokernel of $f \in \mathrm{Hom}_{\mathcal{A}}(x, y)$, the object c is, up to isomorphism, the unique object that satisfies the following universal property: for every object z and every arrow $g\colon y \to z$ such that $gf = 0$, there exists a unique arrow $h\colon c \to z$ such that $h\pi = g$.

b) Show that every cokernel in $\mathcal{A}$ is epi.

Remark 7.21. Let $\mathcal{A}$ be an additive category, and let f be any arrow such that $\ker\operatorname{coker} f$ and $\operatorname{coker}\ker f$ exist. Since $f \circ \ker f = 0$, then by the universal property of the cokernel f factors uniquely through $\operatorname{coker}(\ker f)$, say by $\operatorname{coker}\ker f \xrightarrow{g} y$. Now

$$\operatorname{coker} f \circ g \circ (\operatorname{coker}\ker f) = \operatorname{coker} f \circ f = 0.$$

By Theorem 7.32, $\operatorname{coker}\ker f$ is an epi, and thus we must have $\operatorname{coker} f \circ g = 0$. Then g factors uniquely through $\ker\operatorname{coker} f$, so we get a unique arrow such that

$$\begin{array}{ccccccc} \ker f & \longrightarrow & x & \xrightarrow{\quad f \quad} & y & \longrightarrow & \operatorname{coker} f \\ & & \downarrow & & \uparrow & & \\ & & \operatorname{coker}\ker f & \xrightarrow[\exists!]{} & \ker\operatorname{coker} f & & \end{array}$$

commutes.

Example 7.22. Let's see what this factoring looks like in the more familiar example of R-**Mod**. Given an R-module homomorphism $f\colon M \to N$, we get a commutative diagram

$$\begin{array}{ccccccc} \ker f & \longrightarrow & M & \xrightarrow{\quad f \quad} & N & \longrightarrow & \operatorname{coker} f \cong N/\operatorname{im} f \\ & & \downarrow & & \uparrow & & \\ & & M/\ker f & \xrightarrow[\psi]{} & \operatorname{im} f & & \end{array}$$

The map ψ is the isomorphism given by the First Isomorphism Theorem.

Definition 7.23. An **abelian category** is an additive category $\mathcal{A}$ such that

- The category $\mathcal{A}$ contains all kernels and cokernels of arrows in $\mathcal{A}$.
- Every mono is a kernel of its cokernel.
- Every epi is the cokernel of its kernel.
- For every f, the canonical arrow $\operatorname{coker}\ker f \longrightarrow \ker\operatorname{coker} f$ is an isomorphism.

Ultimately, an abelian category is one that has just enough structure so that we can extend many of the desired properties of R-**Mod**. In particular, we will see that we can define complexes and their homology in any abelian category, and that the Snake Lemma and the long exact sequence in homology hold.

Remark 7.24. Let $\mathcal{A}$ be an abelian category, and f any arrow. As described in Theorem 7.21, we have a commutative diagram

$$\begin{array}{ccccccc} \ker f & \longrightarrow & x & \xrightarrow{\quad f \quad} & y & \longrightarrow & \operatorname{coker} f \\ & & {\scriptstyle p}\downarrow & & \uparrow{\scriptstyle i} & & \\ & & \operatorname{coker}\ker f & \xrightarrow[\psi]{\cong} & \ker\operatorname{coker} f & & \end{array}$$

where ψ is now assumed to be an iso. Now kernels are mono and cokernels are epi, by Theorem 7.32, and composing an epi (respectively, mono) with an iso gives us an epi (respectively, mono). Therefore, we can factor f as a composition mono $\circ$ epi.

Example 7.25. In R-**Mod**, this factorization is just the factoring through the image of the homomorphism: any R-module homomorphism $f\colon M \to N$ factors as

$$M \twoheadrightarrow \operatorname{im} f \hookrightarrow N.$$

Definition 7.26. Let $\mathcal{A}$ be an abelian category, and consider an arrow $x \xrightarrow{f} y$. The **image** of f is $\operatorname{im} f := \ker(\operatorname{coker} f)$.

Following Theorem 7.24, the source of $\operatorname{im} f = \ker\operatorname{coker} f$ is the unique (up to unique isomorphism) object such that f factors as

$$x \xrightarrow{\text{epi}} \operatorname{im} f \xrightarrow{\text{mono}} y\ .$$

Exercise 7.27. Let $\mathcal{A}$ be an abelian category.

a) Show that f is a mono if and only if $\ker f = 0$.

b) Show that f is an epi if and only if $\operatorname{coker} f = 0$.

Remark 7.28. If $\mathcal{A}$ is an abelian category, its opposite category $\mathcal{A}^{\mathrm{op}}$ is also abelian. This is just a consequence of the fact that all the requirements to be an abelian category come together with the dual requirements, so everything automatically dualizes well.

Example 7.29. Here are some examples and nonexamples of abelian categories.

a) The category R-**Mod** is an abelian category.

b) The category of free R-modules is additive but not abelian, as kernels and cokernels do not exist in general.

c) The category of finitely generated R-modules is abelian if and only if R is noetherian, which is exactly the condition we need to guarantee the existence of kernels and cokernels. For a general (nonoetherian) ring R, the category of noetherian R-modules is abelian.

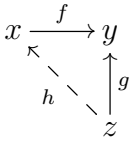
d) The category of Hilbert spaces with continuous linear functions is an additive category. The monos are injective linear maps, and the epis are maps with dense image. The kernels are the usual kernels, while the cokernel of $f : X \longrightarrow Y$ is given by the orthogonal projection $Y \longrightarrow \overline{f(X)}^{\perp}$.

However, we claim that this is not an abelian category, since a mono might not be the kernel of its cokernel. Indeed, if $X \hookrightarrow Y$ is a dense inclusion that is not surjective, then this mono is not the kernel of its cokernel: its cokernel is

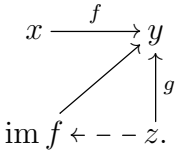
$$Y \longrightarrow \overline{f(X)}^{\perp},$$

but $\overline{f(X)} = Y$ and thus $\overline{f(X)}^{\perp} = \{0\}$, so $\ker\operatorname{coker}(f) = Y$, while $f(X) \neq Y$.

Remark 7.30. Suppose that g factors through f, meaning that there exists h such that



commutes. Then $(\operatorname{coker} f)\circ g = (\operatorname{coker} f)\circ f\circ h = 0$, so g factors through $\ker(\operatorname{coker} f) = \operatorname{im} f$, meaning we have another commutative diagram



Exercise 7.31. Show that the kernel of $x \xrightarrow{0} y$ is the identity arrow 1_x, its cokernel is the identity arrow 1_y, and $\operatorname{im}(x \xrightarrow{0} y) = 0$.

Exercise 7.32. Let $\mathcal{A}$ be an abelian category, g an epi, and f a mono. Then $\ker(fg) = \ker g$, $\operatorname{coker}(fg) = \operatorname{coker} f$, and $\operatorname{im}(fg) = \operatorname{im} f = f$.

7.2 Complexes and homology in an abelian category

Definition 7.33. Let $\mathcal{A}$ be an abelian category. A **chain complex** or simply **complex** (C, ∂), which we sometimes write just a C, is a sequence of objects and arrows

$$\cdots \longrightarrow C_n \xrightarrow{\partial_n} C_{n-1} \longrightarrow \cdots$$

such that $\partial_{n-1}\partial_n = 0$ for all n. A map of complexes $f : C \longrightarrow D$ between two chain complexes is a sequence of arrows f_n such that the diagram

$$\begin{array}{ccccccc} \cdots & \longrightarrow & C_n & \xrightarrow{\partial_n} & C_{n-1} & \longrightarrow & \cdots \\ & & \downarrow{\scriptstyle f_n} & & \downarrow{\scriptstyle f_{n-1}} & & \\ \cdots & \longrightarrow & D_n & \xrightarrow[\partial_n]{} & D_{n-1} & \longrightarrow & \cdots \end{array}$$

commutes. The **category of (chain) complexes over** $\mathcal{A}$, denoted $\operatorname{Ch}(\mathcal{A})$, is the category that has objects all chain complexes in $\mathcal{A}$ and arrows all the chain complex maps.

Lemma 7.34. *If $\mathcal{A}$ is an abelian category, then $\operatorname{Ch}(\mathcal{A})$ is also an abelian category.*

Proof sketch. First, note that $\operatorname{Ch}(\mathcal{A})$ is a preadditive category: given two maps of complexes f and g, $f + g$ is obtained degreewise, by taking

$$(f + g)_n := f_n + g_n.$$

The facts that $\mathrm{Hom}_{\mathrm{Ch}(\mathcal{A})}(C, D)$ is an abelian group and that composition is bilinear follow from the analogous facts in $\mathcal{A}$ (exercise). The zero object is the zero complex, which has the zero object in $\mathcal{A}$ in each degree. Given two complexes C and D, their product is taken degreewise:

$$C \times D = \cdots \longrightarrow C_n \times D_n \xrightarrow{\partial_n^C \times \partial_n^D} C_{n-1} \times D_{n-1} \longrightarrow \cdots$$

and each of the projection maps in each degree assemble to make a map of complexes. So $\mathrm{Ch}(\mathcal{A})$ is an additive category.

Let $f : C \to D$ be a map of complexes. We need to show that both the both the kernel and cokernel of f exist. The universal property of $\ker \partial_n$ gives us a unique arrow δ_{n+1} such that

$$\begin{array}{ccccc} \ker f_{n+1} & \longrightarrow & C_{n+1} & \xrightarrow{f_{n+1}} & D_{n+1} \\ \downarrow{\scriptstyle \delta_{n+1}} & & \downarrow{\scriptstyle \partial_{n+1}} & & \downarrow{\scriptstyle \partial_{n+1}} \\ \ker f_n & \longrightarrow & C_n & \xrightarrow[f_n]{} & D_{n-1} \end{array}$$

commutes. The commutativity of $\begin{array}{ccc} \ker f_{n+1} & \longrightarrow & C_{n+1} \\ \downarrow{\scriptstyle \delta_{n+1}} & & \downarrow{\scriptstyle \partial_{n+1}} \\ \ker f_n & \longrightarrow & C_n \end{array}$ together with $\partial_n \partial_{n+1} = 0$ and the

$$\begin{array}{ccc} \ker f_n & \longrightarrow & C_n \\ \downarrow{\scriptstyle \delta_n} & & \downarrow{\scriptstyle \partial_n} \\ \ker f_{n-1} & \longrightarrow & C_{n-1} \end{array}$$

fact that $\ker f_{n-1}$ is a mono imply that $\delta_n \delta_{n+1} = 0$. Finally, we conclude that

$$\cdots \longrightarrow \ker f_n \xrightarrow{\delta_n} \ker f_{n-1} \longrightarrow \cdots$$

is a complex in $\mathrm{Ch}(\mathcal{A})$, and the canonical maps $\ker f_n \to C_n$ assemble into a map of complexes. One can check that the universal property of the kernels $\ker f_n$ forces this complex we just constructed to be $\ker f$. In particular, $\mathrm{Ch}(\mathcal{A})$ has all kernels. Similarly, we construct cokernels in $\mathrm{Ch}(\mathcal{A})$, building on the fact that $\mathcal{A}$ has all cokernels.

Finally, it remains to show that every mono is the kernel of its cokernel and every epi is the cokernel of its kernel. This boils down to the fact that f is a mono if and only if all the f_n are monos, and dually that f is an epi if and only if all the f_n are epis. The conclusion will then follow from our construction of kernels and cokernels and the fact that $\mathcal{A}$ is abelian. Our claim follows from Theorem 7.27 and the fact that $f = 0$ if and only if all $f_n = 0$. □

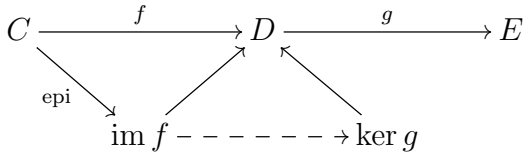
Definition 7.35. Let $\mathcal{A}$ be an abelian category. For each C in $\mathrm{Ch}(\mathcal{A})$, we define its cycles $Z_n(C)$ and boundaries $B_n(C)$ by

$$Z_n(C) := \text{source} \ker \partial_n \quad \text{and} \quad B_n(C) := \text{source} \operatorname{im} \partial_{n+1}.$$

Remark 7.36. Let $\mathcal{A}$ be an abelian category, and $C \xrightarrow{f} D \xrightarrow{g} E$ be arrows in $\mathcal{A}$ such that $gf = 0$. By Theorem 7.24, we can factor f as an epi followed by $\operatorname{im} f$.

$$\begin{array}{ccccc} C & & \xrightarrow{\quad f \quad} & & D \xrightarrow{\quad g \quad} E \\ & \searrow^{\text{epi}} & & \nearrow & \\ & & \operatorname{im} f & & \end{array}$$

Since $g \circ \operatorname{im} f \circ \operatorname{epi} = gf = 0$, we must have $g \circ \operatorname{im} f = 0$, so $\operatorname{im} f$ factors uniquely through $\ker g$. Most importantly, there is a canonical arrow $\operatorname{im} f \longrightarrow \ker g$.



Definition 7.37. Let $\mathcal{A}$ be an abelian category. A sequence of arrows $C \xrightarrow{f} D \xrightarrow{g} E$ in $\mathcal{A}$ is **exact** if $gf = 0$ and $\ker g = \operatorname{im} f$.

Remark 7.38. In our definition of exact sequence, we really mean that the canonical arrow $\operatorname{im} f \longrightarrow \ker g$ we described in Theorem 7.36 is an isomorphism. But notice that is equivalent to saying that the arrow $\operatorname{im} f$ is a kernel for g, and $\ker g$ is an image for f, hence the equality we wrote above, which is a more compact way of saying this.

This immediately generalizes to define an exact sequence, and once again a short exact sequence is one of the form

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0.$$

Exercise 7.39. Show that $0 \longrightarrow A \xrightarrow{f} B$ is exact if and only if f is a mono, and $B \xrightarrow{g} C \longrightarrow 0$ is exact if and only if g is an epi. Moreover,

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

is a short exact sequence if and only if

- f is a mono.
- g is an epi.
- $f = \operatorname{im} f = \ker g$.
- $\operatorname{coker} f = g$.

Remark 7.40. Let $\mathcal{A}$ be an abelian category and (C, ∂) be a complex in $\operatorname{Ch}(R)$. Since $\partial_n \partial_{n+1} = 0$ for all n, we get a canonical arrow $B_n(C) \longrightarrow Z_n(C)$ for each n.

Exercise 7.41. Given an additive category $\mathcal{A}$, B_n and Z_n are additive functors $\operatorname{Ch}(\mathcal{A}) \longrightarrow \mathcal{A}$. In particular, an arrow $C \xrightarrow{f} D$ induces arrows $Z_n(C) \xrightarrow{Z_n(f)} Z_n(D)$ and $B_n(C) \xrightarrow{B_n(f)} B_n(D)$.

Definition 7.42. Let $\mathcal{A}$ be an abelian category and (C, ∂) a complex in $\operatorname{Ch}(R)$. The nth homology of C is the object

$$\mathrm{H}_n(C) := \text{ target of } \operatorname{coker}(B_n(C) \longrightarrow Z_n(C)),$$

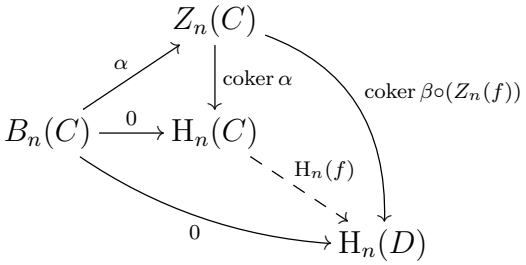
where $B_n(C) \longrightarrow Z_n(C)$ is the canonical arrow we described in Theorem 7.36.

In fact, the nth homology is an additive functor $\mathrm{H}_n \colon \operatorname{Ch}(\mathcal{A}) \longrightarrow \mathcal{A}$. But to see that, we first need to make sense of what homology does to maps of complexes.

Let $\mathcal{A}$ be an abelian category and $C \xrightarrow{f} D$ a map of complexes in $\mathrm{Ch}(\mathcal{A})$. Fix an integer n. We get induced arrows $B_n(f)$ and $Z_n(f)$, since B_n and Z_n are additive functors. This gives us a commutative diagram

$$\begin{array}{ccccc} B_n(C) & \xrightarrow{\alpha} & Z_n(C) & \longrightarrow & \operatorname{coker}\alpha \\ {\scriptstyle B_n(f)}\downarrow & & \downarrow{\scriptstyle Z_n(f)} & & \\ B_n(C) & \xrightarrow[\beta]{} & Z_n(C) & \longrightarrow & \operatorname{coker}\beta \end{array}$$

where α and β are the canonical arrows. To construct $\mathrm{H}_n(f)$, we claim that there is a unique arrow $\operatorname{coker}\alpha \longrightarrow \operatorname{coker}\beta$ making the diagram commute. This is all explained in the commutative diagram

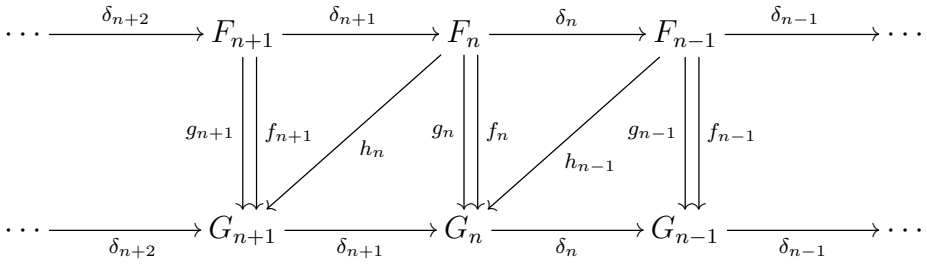


where $\operatorname{coker}\beta \circ (Z_n(f)) \circ \alpha = \operatorname{coker}\beta \circ \beta \circ B_n(f) = 0$, which gives us a unique factorization $\mathrm{H}_n(f)$ through $\operatorname{coker}\alpha$.

Exercise 7.43. Given any abelian category $\mathcal{A}$, H_n is an additive functor $\mathrm{Ch}(\mathcal{A}) \longrightarrow \mathcal{A}$.

Similarly, we can define homotopies.

Definition 7.44. Let $\mathcal{A}$ be an abelian category and $f, g\colon F \longrightarrow G$ be maps of complexes in $\mathrm{Ch}(\mathcal{A})$. A **homotopy**, sometimes referred to as a **chain homotopy**, between f and g is a sequence of arrows $h_n\colon F_n \longrightarrow G_{n+1}$



such that

$$\delta_{n+1}h_n + h_{n-1}\delta_n = f_n - g_n$$

for all n. If there exists a homotopy between f and g, we say that f and g are **homotopic**, and write $f \simeq g$. If f is homotopic to the zero map, we say f is **null-homotopic**. If $f\colon F \to G$ and $g\colon G \to F$ are maps of complexes such that fg is homotopic to the identity arrow 1_G and gf is homotopic to the identity arrow 1_F, we say that f and g are **homotopy equivalences** and F and G are **homotopy equivalent**.

Exercise 7.45. Homotopy is an equivalence relation in $\mathrm{Ch}(\mathcal{A})$.

Exercise 7.46. Let $\mathcal{A}$ be an abelian category. Homotopic maps of complexes in $\mathrm{Ch}(\mathcal{A})$ induce the same map on homology.

Remark 7.47. Let F be an additive functor between abelian categories. Then F must send complexes to complexes, and it induces a functor $\mathrm{Ch}(\mathcal{A}) \to \mathrm{Ch}(\mathcal{A})$, which we also call F. Now if h is a homotopy between two maps of complexes, F preserves the identities $\delta_{n+1}h_n + h_{n-1}\delta_n = f_n - g_n$ for all n, so $F(h)$ is a homotopy between $F(f)$ and $F(g)$.

Definition 7.48. Let $\mathcal{A}$ be an abelian category. A map of complexes $f\colon A \to B$ in $\mathrm{Ch}(\mathcal{A})$ is a **quasi-isomorphism** if $\mathrm{H}_n(f)$ is an isomorphism for all n.

Finally, we set up some notation we will use later.

Definition 7.49. We will denote the full subcategory of $\mathrm{Ch}(\mathcal{A})$ of complexes C such that $C_n = 0$ for all $n < k$ by $\mathrm{Ch}_{\geqslant k}(\mathcal{A})$.

7.3 Functors

Definition 7.50. Let $\mathcal{A}$ be an abelian category. A subcategory $\mathcal{B}$ of $\mathcal{A}$ is an **abelian subcategory** of $\mathcal{A}$ if $\mathcal{B}$ is abelian and the inclusion $\mathcal{B} \subseteq \mathcal{A}$ is an exact functor.

Exercise 7.51. Let $\mathcal{B}$ be a full subcategory of the abelian category $\mathcal{A}$. Show that:

a) $\mathcal{B}$ is an additive category if and only if $\mathcal{B}$ contains 0 and is closed under finite coproducts.

b) $\mathcal{B}$ is an abelian subcategory if and only if $\mathcal{B}$ is additive and closed under kernels and cokernels.

Definition 7.52. Let $T\colon \mathcal{A} \longrightarrow \mathcal{B}$ be an additive covariant functor between abelian categories. We say T is **left exact** if it takes every exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

to the exact sequence

$$0 \longrightarrow T(A) \xrightarrow{T(f)} T(B) \xrightarrow{T(g)} T(C),$$

and **right exact** if it takes every exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

to the exact sequence

$$T(A) \xrightarrow{T(f)} T(B) \xrightarrow{T(g)} T(C) \longrightarrow 0\,.$$

Finally, T is an **exact functor** if it preserves short exact sequences, meaning every short exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

is taken to the short exact sequence

$$0 \longrightarrow T(A) \xrightarrow{T(f)} T(B) \xrightarrow{T(g)} T(C) \longrightarrow 0\,.$$

A contravariant additive functor $T\colon \mathcal{A} \longrightarrow \mathcal{B}$ between abelian categories is **left exact** if it takes every short exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

to the exact sequence

$$0 \longrightarrow T(C) \xrightarrow{T(g)} T(B) \xrightarrow{T(f)} T(A),$$

and **right exact** if it takes every exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

to the exact sequence

$$T(C) \xrightarrow{T(g)} T(B) \xrightarrow{T(f)} T(A) \longrightarrow 0\,.$$

Finally, T is an **exact functor** if it preserves short exact sequences, meaning every short exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

is taken to the short exact sequence

$$0 \longrightarrow T(C) \xrightarrow{T(g)} T(B) \xrightarrow{T(f)} T(A) \longrightarrow 0\,.$$

Theorem 7.53. *Let $\mathcal{A}$ be an abelian category, and fix an object x in $\mathcal{A}$. The functors*

$$\begin{array}{ccc} \mathcal{A} \longrightarrow \boldsymbol{Ab} & & \mathcal{A} \longrightarrow \boldsymbol{Ab} \\ y \longmapsto \mathrm{Hom}_{\mathcal{A}}(x, y) & \textit{and} & y \longmapsto \mathrm{Hom}_{\mathcal{A}}(y, x) \end{array}$$

are left exact.

Proof. We will show that $\mathrm{Hom}_{\mathcal{A}}(x, -)$ is left exact. Notice that the contravariant functor $\mathrm{Hom}_{\mathcal{A}}(-, x)$ can be viewed as the covariant functor $\mathrm{Hom}_{\mathcal{A}^{\mathrm{op}}}(x, -)$. Since $\mathcal{A}^{\mathrm{op}}$ is also an abelian category, it will then follow that $\mathrm{Hom}_{\mathcal{A}^{\mathrm{op}}}(x, -)$ is also left exact, or equivalently, that $\mathrm{Hom}_{\mathcal{A}}(-, x)$ is left exact.

So let

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

be an exact sequence in $\mathrm{Ch}(\mathcal{A})$. We want to show that

$$0 \longrightarrow \mathrm{Hom}_{\mathcal{A}}(x, A) \xrightarrow{f_*} \mathrm{Hom}_{\mathcal{A}}(x, B) \xrightarrow{g_*} \mathrm{Hom}_{\mathcal{A}}(x, C)$$

is exact, and notice this last complex lives in the category of abelian groups.

We have three things to show:

- Exactness at A is equivalent to f being a mono. By assumption, f is a mono, so $f_*(h) = fh$ is injective.
- Since $gf = 0$, so is $g_*f_* = (gf)_*$.
- We want to show that $\ker g_* = \operatorname{im} f_*$, and these are now maps of abelian groups. So we need to show that every $h \in \mathrm{Hom}_{\mathcal{A}}(x, C)$ such that $gh = 0$ factors uniquely through f, meaning $h = \operatorname{im} f_*$. Our assumption that the original sequence is exact implies that $f = \operatorname{im} f = \ker g$. The universal of property of the kernel gives us that whenever $gh = 0$, h must factor through $\ker g = f = \operatorname{im} f$. □

Exercise 7.54. Let I be any small category. Show that if $\mathcal{A}$ is an abelian category, then so is the category $\mathcal{A}^I$ of functors $I \longrightarrow \mathcal{A}$.

We are now ready for the abelian category version of the Yoneda Lemma; this turns out to be a very useful result.

Theorem 7.55 (Yoneda Embedding for abelian categories)**.** *Let $\mathcal{A}$ be an abelian category. Recall that $\boldsymbol{Ab}^{\mathcal{A}^{op}}$ denotes the category of contravariant functors $\mathcal{A} \to \boldsymbol{Ab}$. The covariant functor*

$$\begin{aligned} \mathcal{A} &\longrightarrow \boldsymbol{Ab}^{\mathcal{A}^{op}} \\ x &\longrightarrow \mathrm{Hom}_{\mathcal{A}}(-, x) \end{aligned}$$

is an embedding into a full subcategory. Moreover, this functor reflects exactness, meaning that if

$$\mathrm{Hom}_{\mathcal{A}}(-, x) \longrightarrow \mathrm{Hom}_{\mathcal{A}}(-, y) \longrightarrow \mathrm{Hom}_{\mathcal{A}}(-, z)$$

is exact, then

$$x \longrightarrow y \longrightarrow z$$

must also be exact.

Proof. First, our functor is injective on objects because our axioms for a category include the assumption that the Hom-sets are all disjoint. Moreover, by the usual version of the Yoneda Lemma the assignment

$$\begin{aligned} \mathrm{Nat}(\mathrm{Hom}_{\mathcal{A}}(-, x), \mathrm{Hom}_{\mathcal{A}}(-, y)) &\longrightarrow \mathrm{Hom}_{\mathcal{A}}(x, y) \\ \eta &\longmapsto \eta_x(1_x) \end{aligned}$$

is a natural bijection. In particular, our functor is indeed full and faithful.

To show that the functor reflects exactness, suppose that

$$\mathrm{Hom}_{\mathcal{A}}(-, x) \xrightarrow{f_*} \mathrm{Hom}_{\mathcal{A}}(-, y) \xrightarrow{g_*} \mathrm{Hom}_{\mathcal{A}}(-, z)$$

is exact. Then $g_*f_* = 0$, so $gf = g_*f_*(1_x) = 0$.

It remains to show that $\ker g = \operatorname{im} f$. Let ψ be the canonical arrow $\operatorname{im} f \longrightarrow \ker g$. The exactness of

$$\operatorname{Hom}_{\mathcal{A}}(-, x) \xrightarrow{f_*} \operatorname{Hom}_{\mathcal{A}}(-, y) \xrightarrow{g_*} \operatorname{Hom}_{\mathcal{A}}(-, z)$$

together with the fact that $g_*(\ker g) = 0$ imply that $\ker g$ factors through f. By Theorem 7.30, $\ker g$ must also factor through $\operatorname{im} f$, say by φ. The universal property of the kernels $\ker g$ and $\operatorname{im} f$ will give us that ψ and φ are inverse isos. □

When $\mathcal{A} = R$-**Mod**, the proof can be simplified: the exactness of

$$\operatorname{Hom}_R(R, A) \xrightarrow{f_*} \operatorname{Hom}_{\mathcal{A}}(R, B) \xrightarrow{g_*} \operatorname{Hom}_{\mathcal{A}}(R, C)$$

together with the natural isomorphism between $\operatorname{Hom}_R(R, -)$ and the identity functor give us that

$$A \xrightarrow{f} B \xrightarrow{g} C$$

is exact.

Here is a fun and very useful application of Theorem 7.55.

Corollary 7.56. *Let (L, R) be an adjoint pair of additive functors $\mathcal{A} \underset{R}{\overset{L}{\rightleftarrows}} \mathcal{B}$ between abelian categories. Then L is right exact, and R is left exact.*

Proof. Consider a short exact sequence

$$0 \longrightarrow x \longrightarrow y \longrightarrow z \longrightarrow 0$$

in $\mathcal{B}$, and let w be an object in $\mathcal{A}$. The adjointness of the pair (L, R) gives us a commutative diagram

$$\begin{array}{ccccccc}
0 & \longrightarrow & \operatorname{Hom}_{\mathcal{A}}(w, Rx) & \longrightarrow & \operatorname{Hom}_{\mathcal{A}}(w, Ry) & \longrightarrow & \operatorname{Hom}_{\mathcal{A}}(w, Rz) \\
 & & \downarrow & & \downarrow & & \downarrow \\
0 & \longrightarrow & \operatorname{Hom}_{\mathcal{B}}(Lw, x) & \longrightarrow & \operatorname{Hom}_{\mathcal{B}}(Lw, y) & \longrightarrow & \operatorname{Hom}_{\mathcal{B}}(Lw, z)
\end{array}$$

where the vertical maps are bijections of sets. For every w in $\mathcal{A}$, $\operatorname{Hom}_{\mathcal{B}}(Lw, -)$ is left exact, by Theorem 7.53, so the bottom row of the diagram above is exact. We claim this implies that the top row must also be exact. Our vertical maps are a priori only bijection on sets, but it is easy to see that these natural bijections restrict to a bijection between the images of each pair of corresponding maps. Moreover, for any objects A and B, the natural bijection $\operatorname{Hom}_{\mathcal{A}}(A, RB) \cong \operatorname{Hom}_{\mathcal{A}}(LA, B)$ must always send 0 to 0, since

$$\begin{array}{ccc}
0 = \operatorname{Hom}_{\mathcal{A}}(A, R(0)) & \longrightarrow & \operatorname{Hom}_{\mathcal{A}}(A, R(B)) \\
\downarrow & & \downarrow \\
0 = \operatorname{Hom}_{\mathcal{A}}(L(A), 0) & \longrightarrow & \operatorname{Hom}_{\mathcal{A}}(LA, B)
\end{array}$$

commutes. It is then routine to check that our bijections also restrict to bijections between the kernels of each pair of corresponding maps. The exactness of the bottom row then induces exactness of the top row. By Theorem 7.55, Hom reflects exactness, and we conclude that

$$0 \longrightarrow Rx \longrightarrow Ry \longrightarrow Rz$$

must also be exact. Thus R is a left exact functor.

Finally, by Theorem 7.28, $\mathcal{A}^{\mathrm{op}}$ and $\mathcal{B}^{\mathrm{op}}$ are both abelian categories. Consider the opposite functors L^{op} and R^{op}. Notice that L^{op} is the right adjoint to R^{op}, so L^{op} must be left exact. Therefore, L must be right exact. □

This is possibly the first time we have encountered a proof that truly used duality in an essential and interesting way. In the case where $\mathcal{A} = \mathcal{B} = R$-**Mod**, the fact that R is left exact can be obtained using only methods from R-**Mod**; but the statement about L used the fact that $\mathcal{A}^{\mathrm{op}}$ is an abelian category, while the opposite category of R-**Mod** is not another category of modules.

The Yoneda embedding from Theorem 7.55 is the first piece of the proof of a very important result.

Theorem 7.57 (Freyd-Mitchell embedding theorem)**.** *Let $\mathcal{A}$ be a small abelian category. There exists a ring R, possibly not commutative, and an exact, fully faithful embedding $\mathcal{A} \longrightarrow R$-**Mod**.*

The full details of the proof are rather complicated, and can be found in [Fre03]. Here is a very rough map of the proof. By Theorem 7.55, we already have a fully faithful embedding of $\mathcal{A}$ in $\mathbf{Ab}^{\mathcal{A}^{\mathrm{op}}}$, so it is sufficient to show that there is a fully faithful embedding of $\mathbf{Ab}^{\mathcal{A}^{\mathrm{op}}}$ into some R-**Mod**. The idea is to quotient $\mathbf{Ab}^{\mathcal{A}^{\mathrm{op}}}$ by an abelian subcategory L that contains all the kernels and cokernels of the arrows $\mathrm{Hom}_{\mathcal{A}}(-, y) \to \mathrm{Hom}_{\mathcal{A}}(-, z)$ for all epis $y \longrightarrow z$, in such a way that the composite of the embedding in Theorem 7.55 with this quotient remains an embedding. Then one shows that this quotient category has all coproducts and also what is called a projective generator. Roughly speaking, this is a projective object P such that for every object M there exists an arrow $P \to M$. Then one shows that this implies that this category is equivalent to a full abelian subcategory of R-**Mod** for some R.

Most of the theorems we have proved about R-**Mod** extend to any abelian category. Some of those theorems can in fact be deduced from the fact that they are true over R-**Mod**.

Theorem 7.58 (Snake Lemma)**.** *Consider an abelian category $\mathcal{A}$ and a commutative diagram*

$$\begin{array}{ccccccccc} & & A' & \xrightarrow{i'} & B' & \xrightarrow{p'} & C' & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle h} & & \\ 0 & \longrightarrow & A & \xrightarrow[i]{} & B & \xrightarrow[p]{} & C & & \end{array} .$$

If the rows of the diagram are exact, then there exists an exact sequence

$$\ker f \longrightarrow \ker g \longrightarrow \ker h \xrightarrow{\partial} \operatorname{coker} f \longrightarrow \operatorname{coker} g \longrightarrow \operatorname{coker} h.$$

Theorem 7.59 (Long exact sequence in homology)**.** *Given a short exact sequence in* $\mathrm{Ch}(R)$

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0,$$

there are connecting arrows $\partial : \mathrm{H}_n(C) \longrightarrow \mathrm{H}_{n-1}(A)$ such that

$$\cdots \longrightarrow \mathrm{H}_{n+1}(C) \xrightarrow{\partial} \mathrm{H}_n(A) \xrightarrow{\mathrm{H}_n(f)} \mathrm{H}_n(B) \xrightarrow{\mathrm{H}_n(g)} \mathrm{H}_n(C) \xrightarrow{\partial} \mathrm{H}_{n-1}(A) \longrightarrow \cdots$$

is an exact sequence.

Theorem 7.60 (The Five Lemma)**.** *Given an abelian category $\mathcal{A}$, consider the following commutative diagram in $\mathcal{A}$ with exact rows:*

$$\begin{array}{ccccccccc} A' & \longrightarrow & B' & \longrightarrow & C' & \longrightarrow & D' & \longrightarrow & E' \\ \downarrow{\scriptstyle a} & & \downarrow{\scriptstyle b} & & \downarrow{\scriptstyle c} & & \downarrow{\scriptstyle d} & & \downarrow{\scriptstyle e} \\ A & \longrightarrow & B & \longrightarrow & C & \longrightarrow & D & \longrightarrow & E \end{array}$$

If b and d are epi and e is a mono, then c is an epi. If b and d are mono and a is epi, then c is mono.

One can prove these by invoking the Freyd-Mitchell theorem and checking that one can go back and forth with our statements between some small subcategory of $\mathcal{A}$ containing our diagram and all the necessary kernels, cokernels, etc, and some R-**Mod** where that category embeds. Alternatively, one can use what are called *members*, as in [ML98, VIII.4.5]. The theory of members is an attempt to fix the main difficulty when dealing with abelian categories: that the objects and arrows are not just sets and functions, so we can't just talk about *members* of the objects and their images by each arrow.

7.4 Projectives and injectives

A lot of the notions we have studied this semester can be extended to the setting of a general abelian category.

Definition 7.61. Let $\mathcal{A}$ be an abelian category. An object P in $\mathcal{A}$ is **projective** if $\mathrm{Hom}_{\mathcal{A}}(P, -)$ is an exact functor. An object E in $\mathcal{A}$ is **injective** if $\mathrm{Hom}_{\mathcal{A}}(-, E)$ is exact.

This generalizes the notion of projective and injective modules.

Remark 7.62. Let $\mathcal{A}$ be an abelian category. An object P is projective if and only if every arrow $P \longrightarrow Y$ factors through every epi $X \longrightarrow Y$:

$$\begin{array}{ccccc} & & P & & \\ & \swarrow & \downarrow & & \\ X & \longrightarrow & Y & \longrightarrow & 0 \end{array}$$

and an object E is injective if and only if every arrow $X \longrightarrow E$ factors through every mono $X \longrightarrow Y$:

$$\begin{array}{ccccc} & & E & & \\ & & \uparrow & \nwarrow & \\ 0 & \longrightarrow & X & \longrightarrow & Y \end{array}$$

Exercise 7.63. Let $\mathcal{A}$ be an abelian category.

a) Show that $\mathrm{Hom}_{\mathcal{A}}(x \oplus y, z) = \mathrm{Hom}_{\mathcal{A}}(x, z) \oplus \mathrm{Hom}_{\mathcal{A}}(y, z)$.

b) Show that if P and Q are projective, then so is $P \oplus Q$.

Definition 7.64. An abelian category $\mathcal{A}$ has **enough projectives** if for every object M there exists a projective object P and an epi $P \longrightarrow M$. We say that $\mathcal{A}$ has **enough injectives** if for every object M there exists an injective object E and a mono $M \longrightarrow E$.

Theorem 4.13 and Theorem 4.33 say that R-**Mod** has enough injectives and enough projectives.

Example 7.65. The category of finite abelian groups has no projectives beside 0. In particular, **Ab** does not have enough projectives.

Definition 7.66. Let M be an object in the abelian category $\mathcal{A}$. A **projective resolution** of M is a complex P

$$\cdots \longrightarrow P_2 \longrightarrow P_1 \longrightarrow P_0 \longrightarrow 0$$

where all the P_n are projective, $\mathrm{H}_0(P) = M$, and $\mathrm{H}_n(P) = 0$ for all $n \neq 0$. An **injective resolution** of M is a cochain complex E

$$0 \longrightarrow E^0 \longrightarrow E^1 \longrightarrow \cdots$$

such that every E^n is injective, $\mathrm{H}^n(E) = 0$ for all $n \neq 0$, and $\mathrm{H}^0(E) = M$.

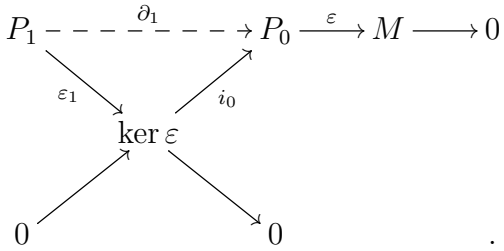
Theorem 7.67. *If $\mathcal{A}$ has enough projectives, every object in $\mathcal{A}$ has a projective resolution. Similarly, if $\mathcal{A}$ has enough injectives, every object in $\mathcal{A}$ has an injective resolution.*

This generalizes Theorem 5.2 in a natural way, and the proof is essentially the same.

Proof. Given be an object M in $\mathcal{A}$, let's construct a projective resolution explicitly. We start by picking an epi $P_0 \xrightarrow{\varepsilon} M$ from a projective P_0. Since ϵ is an epi, it is the cokernel of its kernel, so

$$0 \longrightarrow \ker \varepsilon \longrightarrow P_0 \xrightarrow{\varepsilon} M \longrightarrow 0$$

is a short exact sequence. Now we find an epi $P_1 \xrightarrow{\varepsilon_1} K_0 := \ker \varepsilon$, and set $P_1 \xrightarrow{\partial_1} P_0$ to be the composition



.

We proceed the same way, at each step taking a projective P_n and an epi $\varepsilon_n \colon P_n \longrightarrow \ker \partial_{n-1}$, and setting ∂_{n+1} to be the composition $(\ker \partial_{n-1}) \circ \varepsilon_n$. By construction, $\partial_n = i_{n-1}\varepsilon_n$, where ε_n is an epi and $\ker \partial_{n-1}$ is mono. By Theorem 7.32, $\operatorname{im} \partial_n = i_{n-1} = \ker \partial_{n-1}$. □

We can also characterize injectives in term of split short exact sequences, as we did for modules. In particular, the Splitting Lemma extends to any abelian category.

Definition 7.68. Let $\mathcal{A}$ be an abelian category. A short exact sequence

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

splits if one of the following equivalent conditions hold:

1) There exists an arrow $C \xrightarrow{r} B$ such that $gr = \mathrm{id}_C$.
2) There exists an arrow $B \xrightarrow{s} A$ such that $sf = \mathrm{id}_A$.
3) There exists an isomorphism of complexes between our sequence and

$$0 \longrightarrow A \longrightarrow A \oplus C \longrightarrow C \longrightarrow 0$$

where the arrows are the canonical arrows that come with the (co)product $A \oplus C$.

Theorem 7.69. *Let $\mathcal{A}$ be an abelian category. Every short exact sequence*

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{f} C \longrightarrow 0$$

where A is injective or C is projective splits.

The proofs are exactly the same as in the case of R-**Mod**, Theorem 4.6 and Theorem 4.34.

Proof. If C is projective, there exists h such that

$$\begin{array}{ccccc} & & & & C \\ & & & \swarrow^{h} & \| \\ B & \xrightarrow[g]{} & C & \longrightarrow & 0 \end{array}$$

commutes, so $gh = \mathrm{id}_C$ and g is a splitting. If A is injective, there exists h such that

$$\begin{array}{ccccc} & & A & & \\ & & \| & \nwarrow^{h} & \\ 0 & \longrightarrow & A & \xrightarrow[f]{} & B. \end{array}$$

commutes, so $hf = \mathrm{id}_A$, and h is a splitting. □

More generally, we can talk about split exact complexes.

Definition 7.70. A complex C in $\mathrm{Ch}(\mathcal{A})$ is **split** if there are arrows $s_n \colon C_n \longrightarrow C_{n+1}$ such that the differential ∂ satisfies $\partial = \partial s \partial$. A complex is **split exact** if it is both exact and split.

Remark 7.71. A split short exact sequence is precisely a short exact sequence that is a split complex.

Exercise 7.72. Additive functors preserve split complexes, meaning that if C is a split complex, then so is $F(C)$ for any additive functor F. In particular, additive functors preserve split short exact sequences.

Lemma 7.73. *Let $\mathcal{A}$ be an abelian category, (P, ∂) in $\mathrm{Ch}_{\geqslant 0}(\mathcal{A})$ with each P_i projective, $P_0 \xrightarrow{\partial_0} M$ an arrow in $\mathcal{A}$ such that $\partial_0 \partial_1 = 0$ and (Q, δ) a projective resolution of N. Given any $M \xrightarrow{f} N$ in $\mathcal{A}$, there exists a map of complexes $P \xrightarrow{\varphi} Q$ such that*

$$\begin{array}{ccc} P_0 & \xrightarrow{\partial_0} & M \\ {\scriptstyle \varphi_0}\downarrow & & \downarrow{\scriptstyle f} \\ Q_0 & \xrightarrow[\delta_0]{} & N \end{array}$$

commutes, which is unique up to homotopy.

Proof. Since P_0 is projective and δ_0 is an epi, there exists φ_0 such that

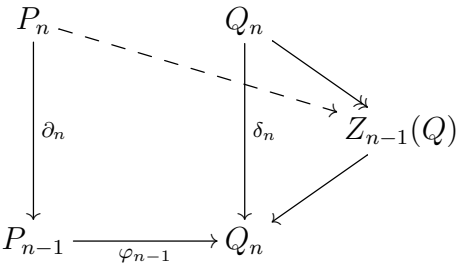
$$\begin{array}{ccccc} P_0 & \xrightarrow{\partial_0} & M & \longrightarrow & 0 \\ {\scriptstyle \varphi_0}\downarrow & & \downarrow{\scriptstyle f} & & \\ Q_0 & \xrightarrow[\delta_0]{} & N & \longrightarrow & 0 \end{array}$$

commutes.

We proceed inductively, assuming we have $\varphi_0, \ldots, \varphi_{n-1}$ with $\varphi_{n-2}\partial_{n-1} = \delta_{n-2}\ \varphi_{n-1}$. Since P_n is projective, there exists φ_n such that

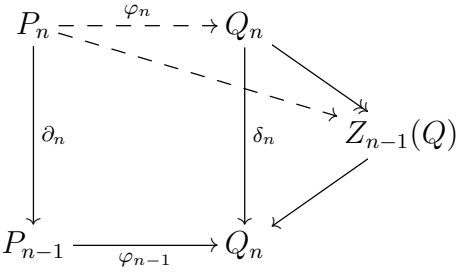
$$\begin{array}{ccccc} P_n & \xrightarrow{\partial_n} & P_{n-1} & \xrightarrow{\partial_{n-1}} & P_{n-2} \\ {\scriptstyle \varphi_n}\downarrow & & \downarrow{\scriptstyle \varphi_{n-1}} & & \downarrow{\scriptstyle \varphi_{n-2}} \\ Q_n & \xrightarrow[\delta_n]{} & Q_n & \xrightarrow[\delta_{n-1}]{} & Q_{n-1} \end{array}$$

commutes. Commutativity gives $\delta_{n-1}\varphi_{n-1}\partial_n = \varphi_{n-2}\partial_{n-1}\partial_n = 0$, so $\varphi_{n-1}\partial_n$ factors through the kernel of δ_{n-1}.



Since Q is a projective resolution of N, the arrow $Q_n \longrightarrow Z_{n-1}(Q)$ above is an epi, so the

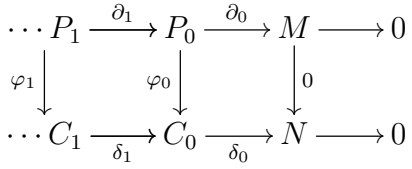
arrow $P_n \longrightarrow Z_{n-1}(Q)$ we just constructed factors through Q_n, giving us φ_n such that



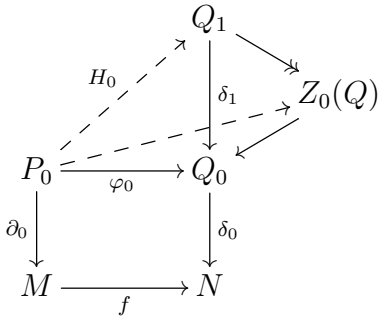
commutes.

Now suppose we are given two such maps of complexes φ and ψ lifting f, say φ and ψ. Note that $\varphi - \psi$ and 0 are two liftings of the 0 map. We are going to show that any map lifting the 0 map $M \longrightarrow N$ must be nullhomotopic, which will then imply that φ and ψ are homotopic as well (essentially via the same homotopy!).

So let $\varphi\colon P \longrightarrow C$ be a map of complexes lifting the 0 map $M \longrightarrow N$.



We will construct a nullhomotopy for φ inductively. Set $h_n = 0$ for all $n < 0$. The commutativity of the rightmost square says that $\delta_0\varphi_0 = 0$, so $\operatorname{im}\varphi_0 \subseteq \ker\delta_0 = \operatorname{im}\delta_1$. Since $\partial_0\varpi_0 = 0$, φ_0 factors through $Z_0(Q)$. But $Q_1 \twoheadrightarrow Z_0(Q)$ is an epi and P_0 is projective, there exists H_0 such that



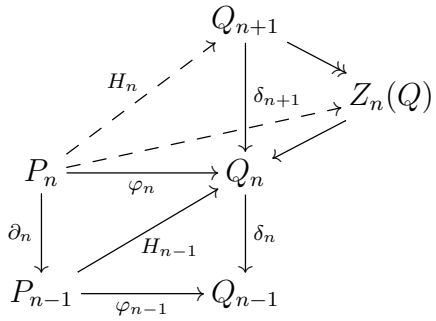
commutes. So H_0 satisfies $\delta_0 H_0 = \varphi_0$. Set $H_{-1} = 0$.

Now suppose we have constructed $H_0, \ldots, H_{n-1}$ such that $\delta_n H_{n-1} + H_{n-2}\partial_{n-1} = \varphi_{n-1}$. Then

$$\begin{aligned}
\delta_n\varphi_n &= \varphi_{n-1}\partial_n && \text{since } \varphi \text{ is a map of complexes}\\
&= (\delta_n H_{n-1} + H_{n-2}\partial_{n-1})\partial_n && \text{by assumption}\\
&= \delta_n H_{n-1}\partial_n + H_{n-2}\partial_{n-1}\partial_n\\
&= \delta_n H_{n-1}\partial_n && \text{since } \partial_{n-1}\partial_n = 0
\end{aligned}$$

so $\delta_n(\varphi_n - H_{n-1}\partial_n) = 0$. Therefore, φ_n factors through $Z_n(Q)$, and since Q is a projective resolution of N, $Q_{n+1} \longrightarrow Z_n(Q)$ is an epi. Therefore, the factorization of $\varphi_n - H_{n-1}\partial_n$

through $Z_n(Q)$ also factors through Q_n, and we end up with an arrow H_n such that



commutes. This H_n must then satisfy $\delta_{n-1}H_n + H_{n-1}\partial_n = \varphi_n$, and ultimately H is a homotopy between φ and 0. □

Theorem 7.74 (Horseshoe Lemma)**.** *Let $\mathcal{A}$ be an abelian category, P be a projective resolution of A, and R be a projective resolution of C. If*

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

is an exact sequence, there exists a projective resolution Q of B and maps of complexes F and G lifting f and g such that

$$0 \longrightarrow P \xrightarrow{F} Q \xrightarrow{G} R \longrightarrow 0$$

is an exact sequence in $\mathrm{Ch}(\mathcal{A})$.

Proof. First, a word on notation: $\oplus$ denotes the coproduct in $\mathcal{A}$, and given arrows $x \xrightarrow{f} z$ and $y \xrightarrow{g} z$, we will write $f \oplus g$ for the unique arrow $x \oplus y \longrightarrow z$ induced by f and g. Moreover, we will denote the differential of P by ∂^P, and the differential of R by ∂^R.

Set $Q_n = P_n \oplus R_n$. Recall that the product and coproduct in $\mathcal{A}$ coincide, by Theorem 7.10, so let $F_n \colon P_n \longrightarrow Q_n$ and $G_n \colon Q_n \longrightarrow R_n$ be the canonical arrows. One can show that in fact we get short exact sequences

$$0 \longrightarrow P_n \xrightarrow{F_n} Q_n \xrightarrow{G_n} R_n \longrightarrow 0$$

for all n. Moreover, Q_n is projective for all n, by Theorem 7.63. We will construct the missing differentials ∂^Q inductively.

Since R_0 is projective and g is an epi, there exists γ such that

$$\begin{array}{ccccccccc} 0 & \longrightarrow & P_0 & \xrightarrow{F_0} & Q_0 & \xrightarrow{G_0} & R_0 & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \partial_0} & & & \swarrow_{\gamma} & \downarrow{\scriptstyle \partial_0} & & \\ 0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow[g]{} & C & \longrightarrow & 0 \end{array}$$

commutes. Set $\partial_0^Q := (f\partial_0^P) \oplus \gamma$. The universal property of the coproduct guarantees that

$$\begin{array}{ccccccccc} 0 & \longrightarrow & P_0 & \xrightarrow{F_0} & Q_0 & \xrightarrow{G_0} & R_0 & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \partial_0} & & \downarrow{\scriptstyle \partial_0^Q} & \swarrow_{\gamma} & \downarrow{\scriptstyle \partial_0} & & \\ 0 & \longrightarrow & A & \xrightarrow[f]{} & B & \xrightarrow[g]{} & C & \longrightarrow & 0 \end{array}$$

commutes. By the Five Lemma, ∂_0^Q is epi. By the Snake Lemma,

$$\ker \partial_0^P \longrightarrow \ker \partial_0^Q \longrightarrow \ker \partial_0^R$$

is exact. We then proceed by induction, and at each step we apply the base case to

$$\begin{array}{ccccccccc} 0 & \longrightarrow & P_{n+1} & \xrightarrow{F_{n+1}} & Q_{n+1} & \xrightarrow{G_{n+1}} & R_n & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \partial^P_{n+1}} & & & & \downarrow{\scriptstyle \partial^R_{n+1}} & & \\ 0 & \longrightarrow & \ker \partial_n^P & \longrightarrow & \ker \partial_n^Q & \longrightarrow & \ker \partial_n^R & \longrightarrow & 0 \end{array}$$

where the vertical arrows are epi because P and R are projective resolutions and thus exact. □

Remark 7.75. By duality, if $\mathcal{A}$ has enough injectives,

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

is exact, and E_A and E_C are injective resolutions for A and C, then there exist an injective resolution E_B of B and a short exact sequence of complexes

$$0 \longrightarrow E_A \longrightarrow E_B \longrightarrow E_C \longrightarrow 0$$

extending the given one.

7.5 Derived functors

We are now ready to define derived functors in full generality. The definitions will match the definitions over R-modules; the one notable addition from R-**Mod** to the general case is the new need to worry about whether the abelian category in question has enough injectives or enough projectives.

Definition 7.76 (Derived functors)**.** Let $\mathcal{A}$ and $\mathcal{B}$ be abelian categories.

Let $F\colon \mathcal{A} \longrightarrow \mathcal{B}$ be a covariant right exact functor. If $\mathcal{A}$ has enough projectives, the **left derived functors** of F are a sequence of functors $L_iF\colon \mathcal{A} \longrightarrow \mathcal{B}$, $i \geqslant 0$, defined as follows:

- For each object A in $\mathcal{A}$, fix a projective resolution P of A, and set
$$L_iF(A) := \mathrm{H}_i(F(P)).$$
- Given an arrow $f\colon A \to B$, fix projective resolutions $P \longrightarrow A$ and $Q \longrightarrow B$, and a map of complexes $\varphi\colon P \to Q$ lifting f. Then
$$L_iF(f) := \mathrm{H}_i(F(\varphi)).$$

Let $F\colon \mathcal{A} \longrightarrow \mathcal{B}$ be a covariant left exact functor. If $\mathcal{A}$ has enough injectives, the **right derived functors** of F are a sequence of functors $R^iF\colon \mathcal{A} \longrightarrow \mathcal{B}$, $i \geqslant 0$, defined as follows:

- For each object A in $\mathcal{A}$, fix an injective resolution E of A, and set

$$R^iF(A) := \mathrm{H}^i(F(E)).$$

- Given an arrow $f\colon A \to B$, fix injective resolutions $A \longrightarrow E$ and $B \longrightarrow I$, and a map of complexes $\varphi\colon P \to Q$ extending f. Then

$$R^iF(f) := \mathrm{H}^i(F(\varphi)).$$

Let $F\colon \mathcal{A} \longrightarrow \mathcal{B}$ be a contravariant left exact functor. If $\mathcal{A}$ has enough projectives, the **right derived functors** of F are a sequence of functors $R^iF\colon \mathcal{A} \longrightarrow \mathcal{B}$, $i \geqslant 0$, defined as follows:

- For each object A in $\mathcal{A}$, fix a projective resolution P of A, and set

$$R^iF(A) := \mathrm{H}^i(F(P)).$$

- Given an arrow $f\colon A \to B$, fix projective resolutions $P \longrightarrow A$ and $Q \longrightarrow B$, and a map of complexes $\varphi\colon P \to Q$ extending f. Then

$$R^iF(f) := \mathrm{H}^i(F(\varphi)).$$

Finally, let $F\colon \mathcal{A} \longrightarrow \mathcal{B}$ be a contravariant right exact functor. If $\mathcal{A}$ has enough injectives, the **left derived functors** of F are a sequence of functors $L_iF\colon \mathcal{A} \longrightarrow \mathcal{B}$, $i \geqslant 0$, defined as follows:

- For each object A in $\mathcal{A}$, fix an injective resolution E of A, and set

$$L_iF(A) := \mathrm{H}_i(F(E)).$$

- Given an arrow $f\colon A \to B$, fix injective resolutions $A \longrightarrow E$ and $B \longrightarrow I$, and a map of complexes $\varphi\colon E \to I$ extending f. Then

$$L_iF(f) := \mathrm{H}_i(F(\varphi)).$$

Remark 7.77. If F is exact, then $\mathrm{H}_i(F(C)) = F(\mathrm{H}_i(C))$, so $L_iF = 0$ for all $i > 0$.

Remark 7.78. If P is projective, then $0 \longrightarrow P \longrightarrow 0$ is a projective resolution of P, and thus $L_iF(P) = 0$ for all $i > 0$. Similarly, if E is injective then $R^iF(E) = 0$.

Proposition 7.79. *Let $\mathcal{A}$ be an abelian category with enough projectives, and F a covariant right exact functor.*

a) $L_iF(A)$ is well-defined up to isomorphism for every object A.

b) $L_iF(f)$ is well-defined for every arrow f.

c) L_iF is an additive functor for each i.

d) $L_0F = F$.

Proof.

a) Let P and Q be projective resolutions of A. Theorem 5.19 gives us maps of complexes $\varphi\colon P \to Q$ and $\psi\colon Q \to P$ such that $\varphi\psi$ is homotopic to 1_Q and $\psi\varphi$ is homotopic to 1_P. Additive functors preserve homotopies, by Theorem 7.47, so $F(\varphi)F(\psi)$ and $F(\psi)F(\varphi)$ are homotopic to the corresponding identity arrows. Homotopic maps induce the same map in homology, by Theorem 7.46. Therefore, $F(\varphi)$ and $F(\psi)$ induce isomorphisms in homology.

b) Fix projective resolutions P and Q of M and N. Any two lifts φ and ψ of $f\colon M \longrightarrow N$ to $P \longrightarrow Q$ are homotopic, by Theorem 7.73. Additive functors preserve homotopies, by Theorem 7.47, so $F(\varphi)$ and $F(\psi)$ are homotopic. Homotopic maps induce the same map in homology, by Theorem 7.46, so $L_iF(\varphi) = L_iF(\psi)$ for each i.

c) Given an arrow f, fix a lift φ of f to projective resolutions of the source and target. Since F is an additive functor, $\mathrm{H}_i(F(\varphi))$ is a homomorphism for each i, and thus $L_iF(f)$ is a homomorphism between the corresponding Hom-groups, which as we've seen is independent of our choice of φ.

d) Let A be any object and P be a projective resolution of A. Since P is right exact, and

$$P_1 \longrightarrow P_0 \longrightarrow A \longrightarrow 0$$

is exact, then so is

$$F(P_1) \longrightarrow F(P_0) \longrightarrow F(A) \longrightarrow 0.$$

We claim that $\mathrm{H}_0(F(P)) = F(A)$. By Theorem 7.31, $\ker(F(P_0) \longrightarrow 0) = 1_{F(P_0)}$, so the canonical arrow $\operatorname{im} F(\partial_1) \longrightarrow F(P_0)$ is precisely the image of $F(\partial_1)$. By exactness of the last sequence we wrote above, $\operatorname{im} F(\partial_1) = \ker(F(P_0) \longrightarrow F(A))$. On the other hand, exactness at $F(A)$ says that $F(P_0) \longrightarrow F(A)$ is an epi, by Theorem 7.39. Every epi is the cokernel of its kernel, so $F(P_0) \longrightarrow F(A)$ is the cokernel of $\operatorname{im} F(\partial_1)$, which we saw was exactly the canonical arrow $B_1(F(P)) \longrightarrow Z_0(F(P))$. Therefore, $\mathrm{H}_0(F(P)) = F(A)$, the target of the cokernel of $B_1(F(P)) \longrightarrow Z_0(F(P))$. □

Exercise 7.80. Let $\mathcal{A}$ be an abelian category with enough injectives, and F a covariant left exact functor.

a) $R^iF(A)$ is well-defined up to isomorphism.

b) $R^iF(f)$ is well-defined for every arrow f.

c) $R^iF(f)$ is an additive functor for every i.

d) $R^0F = F$.

Remark 7.81. If $\mathcal{A}$ is an abelian category with enough injectives, then $\mathcal{A}^{\mathrm{op}}$ is an abelian category with enough projectives. This gives us a relationship between left derived and right derived functors: $R^iF = (L_iF^{\mathrm{op}})^{\mathrm{op}}$.

Theorem 7.82. *Let $T_i : \mathcal{A} \longrightarrow \mathcal{B}$ be a sequence of additive covariant functors between abelian categories, where $\mathcal{A}$ has enough projectives, and $F\colon \mathcal{A} \longrightarrow \mathcal{B}$ a right exact functor. Suppose the following hold:*

(1) For every short exact sequence $0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$ in $\mathcal{A}$, we get a natural long exact sequence

$$\cdots \longrightarrow T_2(C) \longrightarrow T_1(A) \longrightarrow T_1(B) \longrightarrow T_1(C) \longrightarrow T_0(A) \longrightarrow T_0(B) \longrightarrow T_0(C) \longrightarrow 0.$$

(2) T_0 is naturally isomorphic to F.

(3) $T_n(P) = 0$ for every projective object P in $\mathcal{A}$, and all $n \geqslant 1$.

Then T_n is naturally isomorphic to L_nF for all n.

Similarly, suppose $T_i\colon \mathcal{A} \longrightarrow \mathcal{B}$ is a sequence of additive covariant functors, where $\mathcal{A}$ has enough injectives, and $F\colon \mathcal{A} \longrightarrow \mathcal{B}$ a left exact functor such that

a) For every short exact sequence $0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$ in $\mathcal{A}$, we get a long exact sequence

$$0 \longrightarrow T_0(A) \longrightarrow T_0(B) \longrightarrow T_0(C) \longrightarrow T_1(A) \longrightarrow T_1(B) \longrightarrow T_1(C) \longrightarrow \cdots .$$

b) T_0 is naturally isomorphic to F.

c) $T_n(E) = 0$ for every injective object E in $\mathcal{A}$, and all $n \geqslant 1$.

Then T_n is naturally isomorphic to R^nF for all n.

Chapter 8

Spectral Sequences

> It has been suggested that the name "spectral" was given because, like spectres, spectral sequences are terrifying, evil, and dangerous. I have heard no one disagree with this interpretation, which is perhaps not surprising since I just made it up.
>
> (Ravi Vakil, in *Spectral Sequences: friend or foe?*)

Spectral sequences are useful bookkeeping tools for computing, among other things, the homology and cohomology of complicated complexes. Unfortunately, spectral sequences have a bad reputation for being difficult and scary; but continuing Vakil's quote above, "you can use spectral sequences without hesitation or fear, and [...] you shouldn't be frightened when they come up in a seminar". Rotman wisely says that "Of course, the reader must digest these new ideas in order to apply them, but it is worth the effort" [Rot09, page 608].

Spectral sequences were introduced independently by Leray and Lyndon in the 1940s. Leray came up with the idea while he was a prisoner of the nazis during WWII. The primary usages of spectral sequences are in homotopy theory and topology more generally, but there are also applications in commutative algebra and other fields. Due to its topological roots, the subject is best learned with a topological backdrop – which provides many motivating examples – but since we are not assuming any particular topological background, we will give only a very brief introduction to the subject, and from a more algebraic perspective.

Here are some recommended sources to learn more about spectral sequences:

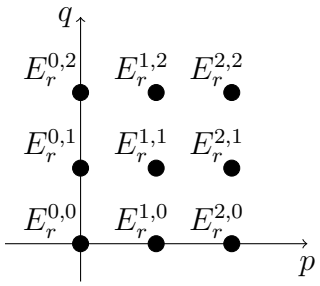
- Rotman's book *An introduction to Homological Algebra* (second edition) [Rot09].
- John McCleary's *A user's guide to spectral sequences* [McC01].
- Ravi Vakil's notes *Spectral Sequences: friend or foe?*.
- Hatcher's additional chapter on spectral sequences, an addition to his *Algebraic Topology* book [Hat02].
- Weibel's book *An introduction to homological algebra* [Wei94].
- Eisenbud's Appendix A3.13 to his book *Commutative algebra with a view towards algebraic geometry* [Eis95].
- Mel Hochster's notes.
- Michael Hutchings notes.

8.1 What is a spectral sequence?

Roughly speaking, a spectral sequence is a book where each page is a plane with a module (or more general, an object in an abelian category) sitting in each point with integer coordinates. To pass the pages, we take the (co)homology of a differential, and attach to the next page a new differential with a different shape.

Definition 8.1. A **cohomological spectral sequence** E consists of the following data:

- A family $E = (E_r^{p,q})$ of R-modules ranging over all integers p, q, r with $r \geqslant 0$. For a fixed r, the collection of modules $E_r = E_r^{p,q}$ is called the rth **page** or **sheet** of the spectral sequence. We think of each page as living in $\mathbb{Z}^2$, and depict the rth page by putting $E_r^{p,q}$ in the point with coordinates (p, q):



- Differentials

$$d_r \colon E_r^{p,q} \longrightarrow E_r^{p+r,q-r+1},$$

  so R-module homomorphisms such that $d_r d_r = 0$, or more precisely,

$$d_r^{p+r,p-r+1} d_r^{p,q} = 0.$$

- Isomorphisms

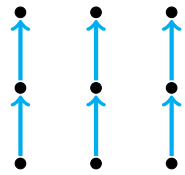
$$E_{r+1}^{p,q} \cong \mathrm{H}^{p,q}(E_r),$$

  meaning

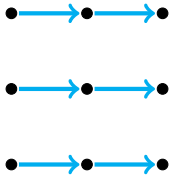
$$E_{r+1}^{p,q} = \frac{\ker\left(E_r^{p,q} \xrightarrow{d_r} E_{p+r,q-r+1}^r \right)}{\operatorname{im}\left(E_{p-r,q+r-1}^r \xrightarrow{d_r} E_{p,q}^r \right)}$$

  for every p, q.

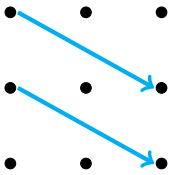
Remark 8.2. The differential in the rth page is a map of degree $(r, -r+1)$. Here are some examples:



0th page



1st page

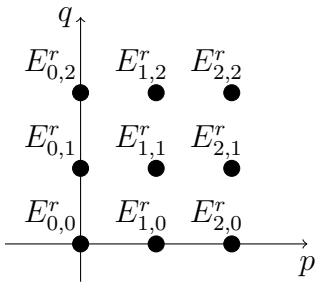


2nd page

We can define spectral sequences more generally over any abelian category. However, most spectral sequences the typical mathematician ever computes are over R-modules (or even just vector spaces!), so we will keep things friendly by sticking to R-modules.

Definition 8.3. A **homological spectral sequence** E is a sequence consists of the following data:

- A family $E = (E^r_{p,q})$ of R-modules ranging over all integers p, q, r with $r \geqslant 0$. For a fixed r, the collection of objects $E_r = E_r^{p,q}$ is called the rth **page** or **sheet** of the spectral sequence. We think of each page as living in $\mathbb{Z}^2$, and depict the rth page by putting $E_r^{p,q}$ in the point with coordinates (p, q):



- Differentials

$$d_r \colon E^r_{p,q} \longrightarrow E^r_{p-r,q+r-1}$$

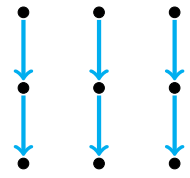
and

- for each r, isomorphisms $E^{r+1} \cong \mathrm{H}(E^r, d^r)$, meaning

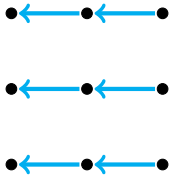
$$E^{r+1}_{p,q} \cong \frac{\ker\left(E^r_{p,q} \xrightarrow{d_r} E^r_{p-r,q+r-1} \right)}{\mathrm{im}\left(E^r_{p+r,q-r+1} \xrightarrow{d_r} E^r_{p,q} \right)}$$

for every p, q.

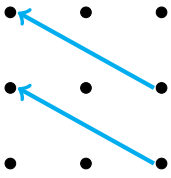
Remark 8.4. The differential in the rth page is a map of degree $(-r, r-1)$.



0th page



1st page



2nd page

8.2 Graded and bigraded modules and their filtrations

We can be a bit more formal and say that a page in a spectral sequence is a differential bigraded module. The goal of this section is to set up some background on the topic of differential bigraded modules.

Definition 8.5. Let R be a ring. A **graded module** over R is a family $M = (M_n)_{n\in\mathbb{Z}}$ of R-modules indexed by $\mathbb{Z}$. We sometimes denote M by $M_\bullet$.

More precisely, these are $\mathbb{Z}$-graded modules. More generally, we have discussed $\mathbb{Z}$-graded modules over any graded ring; in that case, the action of R on M must respect the grading. But here we are not assuming any grading on R, so we have no such conditions. One may reinterpret our definition of graded module as assuming that R is given the trivial grading concentrated in degree 0, so that $R_0 = R$ and $R_i = 0$ otherwise, and thus the condition on the action of R becomes trivial.

Example 8.6. If C is a complex of R-modules, then $(C_n)_{n\in\mathbb{Z}}$ is a graded module. Moreover, its homology $H = (H_n(C))$ also forms a graded module.

Definition 8.7. Given graded modules M and N, a **graded map** of degree d is a family

$$f = (f_n \colon M_n \to N_{n+d})_{n\in\mathbb{Z}}$$

of homomorphisms of R-modules, which we denote by $f \colon M \to N$. We write $\deg(f) = d$ to denote that the degree f is d.

Example 8.8.

1) If C is a complex of R-modules, the differential d is a graded map $d \colon C \to C$ of degree -1. If C is a cochain complex, the differential is a graded map of degree 1.

2) Any map of complexes $f \colon C \to D$ is a graded map of degree 0.

3) A homotopy is a map of degree 1.

Definition 8.9. Let R be a ring. The **category of graded modules** over R has objects all graded modules and arrows all graded maps of graded modules over R.

Definition 8.10. Let M and N be graded modules over a ring R. We say N is a **submodule** of M, and write $N \subseteq M$, if $N_n \subseteq M_n$ for all n. The **quotient** of M by N is the graded module

$$M/N := (M_n/N_n)_n.$$

Definition 8.11. Let $f : M \to N$ be a graded map of degree d between graded modules. The **kernel** of f is the graded module

$$\ker f := (\ker f_n)_{n\in\mathbb{Z}}$$

and the **image** of f is the graded module

$$\operatorname{im} f := (\operatorname{im} f_{n-d})_{n\in\mathbb{Z}}.$$

Definition 8.12. Consider graded maps of graded modules

$$A \xrightarrow{f} B \xrightarrow{g} C.$$

We say this is an **exact sequence** if $\operatorname{im} f = \ker g$.

Remark 8.13. By our definition of kernel and image, $\operatorname{im} f = \ker g$ says that $\operatorname{im} f_{n-d} = \ker g_n$ for all n.

Definition 8.14. Let R be a ring. A **bigraded module** over R is a family $M = (M_{p,q})_{p,q\in\mathbb{Z}}$ of R-modules indexed by $\mathbb{Z} \times \mathbb{Z}$. We sometimes denote M by $M_{\bullet,\bullet}$.

Definition 8.15. Let M and N be bigraded modules over a ring R. A **bigraded map** $f\colon M \to N$ of degree (a, b) is a family of homomorphisms of R-modules

$$f = (f_{p,q}\colon M_{p,q} \to N_{p+a,q+b})_{p,q\in\mathbb{Z}}.$$

We denote the degree of f by $\deg(f) = (a, b)$.

The **kernel** of f is the bigraded module

$$\ker f := (\ker f_{p,q})_{p,q\in\mathbb{Z}}$$

and the image of f is the bigraded module

$$\operatorname{im} f := (\operatorname{im} f_{p-a,q-b})_{p,q\in\mathbb{Z}}.$$

We say a sequence of graded maps

$$A \xrightarrow{f} B \xrightarrow{g} C$$

is **exact** if $\operatorname{im} f = \ker g$.

Definition 8.16. A **differential (bi)graded module** (M, d) over a ring R consists of a (bi)graded module M and a graded map $d\colon M \to M$, which we call the differential, such that $dd = 0$.

We can think of (bi)graded modules as differential (bi)graded modules with zero differential.

Example 8.17. A double complex C with differentials d^h and d^v gives rise to two differential bigraded complexes: (C, d^h) and (C, d^v).

Definition 8.18. Let (M, d) be a differential (bi)graded module. The **homology** of M is the (bi)graded module

$$\mathrm{H}(M, d) = \ker d / \operatorname{im} d.$$

We sometimes shorten this to $\mathrm{H}(M)$.

A **spectral sequence** can now be recast as a sequence (E^r, d^r) of differential bigraded modules such that $E^{r+1} = \mathrm{H}(E^r, d^r)$.

8.3 Filtrations

Definition 8.19. Let M be an R-module. A **filtration** $F^\bullet M$ on M is a family $(F^p M)_{p\in\mathbb{Z}}$ of submodules of M such that either

$$F^p M \subseteq F^{p+1} M$$

for all p, in which case we say the filtration is **increasing** or **ascending**, or

$$F^p M \subseteq F^{p-1} M$$

for all p, in which case we say the filtration is **decreasing** or **descending**. When we do not specify if a filtration is ascending or descending, we will assume by default that it is ascending, though we accept both kinds as filtrations in their own right. The **factors** of $F^\bullet$ are the quotient modules

$$F^p M / F^{p-1} M.$$

One can in fact define filtrations on any abelian category; that requires the notion of a subobject, which we have not yet defined, but it is easy to guess: a **suboject** of an object x in an abelian category $\mathcal{A}$ is a mono with target x. In particular, we can define a filtration on a complex, or more generally on a (differential) (bi)graded module:

Definition 8.20. An ascending **filtration** on a (bi)graded module M is a sequence $F^p M$ of submodules of M such that $F^p M \subseteq F^{p+1} M$. If M is a differential (bi)graded module, we require additionally that the filtration respects the differential, that is, that $d(F^p M) \subseteq F^p M$.

So in particular when C is a complex, we get the following definition:

Definition 8.21. Let C be a complex. An ascending **filtration** $F^\bullet C$ of C is an ascending chain of subcomplexes $F^p C$ of C

$$\cdots \subseteq F^{p-1} C \subseteq F^p C \subseteq F^{p+1} C \subseteq \cdots .$$

We call a complex (C, ∂) with a filtration F a **filtered complex**, and denote it by (C, ∂, F).

As above, one can define a descending filtration; if we do not indicate whether a filtration is ascending or descending, we will always by default assume it is ascending.

Remark 8.22. Let C be a complex and consider an ascending filtration $F_\bullet C$ of C. For each fixed homological degree n, we get an ascending filtration

$$\cdots \subseteq F^{p-1} C_n \subseteq F^p C_n \subseteq F^{p+1} C_n \subseteq \cdots$$

of submodules of C_n. A filtration of C gives us commutative diagrams

$$\begin{array}{ccccccc}
\longrightarrow & C_{n+1} & \longrightarrow & C_n & \longrightarrow & C_{n-1} & \longrightarrow \\
 & \uparrow & & \uparrow & & \uparrow & \\
\longrightarrow & F^p C_{n+1} & \longrightarrow & F^p C_n & \longrightarrow & F^p C_{n-1} & \longrightarrow \\
 & \uparrow & & \uparrow & & \uparrow & \\
\longrightarrow & F^{p-1} C_{n+1} & \longrightarrow & F^{p-1} C_n & \longrightarrow & F^{p-1} C_{n-1} & \longrightarrow
\end{array}$$

where the rows are given by the differential on C and its restrictions to the appropriate modules.

Definition 8.23. Let C be a complex with a filtration $F^\bullet C$, and consider the inclusions $i_p \colon F^p C \to C$. We get an **induced filtration** in homology, as follows: for each n,

$$F^p \operatorname{H}_n(C) := \operatorname{im}(\operatorname{H}_n(F^p C) \to \operatorname{H}_n(C)),$$

giving us a filtration $F^\bullet \operatorname{H}(C)$ on the graded module $\operatorname{H}(C)$.

Definition 8.24. A filtration $F^\bullet M$ of a graded module M is **bounded** if for each n there exists integers $s = s(n)$ and $t = t(n)$ such that

$$F^s M_n = 0 \quad \text{and} \quad F^t M_n = M_n,$$

so that the filtration on M_n can be described by finitely many terms

$$F^s M_n = 0 \subseteq F^{s+1} M_n \subseteq \cdots F^{t-1} M_n \subseteq F^t M_n = M_n.$$

In particular, $F^i M_n = 0$ for all $i < s$ and $F^i M_n = M_n$ for all $i > t$.

Notice, however, that the bounds s and t may depend on n, and in particular there is not necessarily global integers s and t such that

$$F^s M_n = 0 \quad \text{and} \quad F^t M_n = M_n$$

for all n.

Remark 8.25. Suppose that $F^\bullet M$ is a bounded filtration on a complex C. Then the induced filtration in homology is also bounded, with the same bounds (or better), so that for all n there exist s and t such that

$$0 = F^s \operatorname{H}_n(C) = 0 \quad \text{and} \quad F^t \operatorname{H}_n(C) = \operatorname{H}_n(C).$$

Definition 8.26. Let M be a module or a graded module. Given an increasing filtration F of M, its **associated graded module** is the graded module $\operatorname{gr}_F(M)$ given by

$$\operatorname{gr}_F(M) := (F^n M / F^{n-1} M)_{n \in \mathbb{Z}}.$$

Given a decreasing filtration F of an R-module M, its **associated graded module** is the graded module $\operatorname{gr}_F(M)$ given by

$$\operatorname{gr}_F(M) := (F^n M / F^{n+1} M)_{n \in \mathbb{Z}}.$$

Note that the associated graded module depends on the choice of filtration.

Remark 8.27. Let (C, d, F) be a filtered complex. The differential d induces a differential on the associated graded:

$$\begin{aligned} F^p C_n / F^{p-1} C_n &\longrightarrow F^p C_{n-1} / F^{p-1} C_{n-1} \\ a + F^{p-1} C_n &\longmapsto d(a) + F^{p-1} C_{n-1}. \end{aligned}$$

This is well-defined, since our definition of filtered complex requires that the filtration respects the differential: if $a \in F_{p-1} C_n$, then $d(a) \in F^{p-1} C_{n-1}$.

Definition 8.28. We can filter the integers $\mathbb{Z}$ as a $\mathbb{Z}$-module as follows: we set

$$F^n\mathbb{Z} = \mathbb{Z} \quad \text{whenever } n \leqslant 0,$$

and

$$F^n\mathbb{Z} = (2^n) \quad \text{whenever } n \geqslant 0.$$

This give us the decreasing filtration

$$\cdots \supseteq \mathbb{Z} \supseteq \mathbb{Z} \supseteq (2) \supseteq (4) \supseteq (8) \supseteq \cdots .$$

The corresponding associated graded module has

$$\mathrm{gr}_F(\mathbb{Z})_n = 0 \quad \text{whenever } n \leqslant 0$$

and

$$\mathrm{gr}_F(\mathbb{Z})_n = (2^n)/(2^{n+1}) \cong \mathbb{Z}/2 \quad \text{whenever } n \geqslant 0.$$

Remark 8.29 (The associated graded versus the actual module)**.** Suppose someone has filtered the R-module M by $F^\bullet M$, but that we only have access to the associated graded module of this filtration. While the pieces of the associated graded give us information about M, they may not be sufficient to fully compute M. For a simple example, consider an R-module B and a submodule $A \subseteq B$, and the filtration

$$F^pM = \begin{cases} 0 & \text{if } p < 0 \\ A & \text{if } p = 0 \\ B & \text{if } p \geqslant 1. \end{cases}$$

The associated graded is given by

$$\mathrm{gr}_n = \begin{cases} 0 & \text{if } p < 0 \\ A & \text{if } p = 0 \\ B/A & \text{if } p = 1 \\ 0 & \text{if } p > 1. \end{cases}$$

So if we are only given the associated graded, meaning, if all we have access to is A and B/A, then all we know about B is that fits into a short exact sequence

$$0 \longrightarrow A \longrightarrow B \longrightarrow B/A \longrightarrow 0.$$

However, if $\mathrm{Ext}^1_R(B/A, A) \neq 0$, then this does not uniquely determine B. Indeed, it turns out that the isomorphism classes of extensions of A by B/A, meaning modules M that fit into short exact sequences of the form

$$0 \longrightarrow A \longrightarrow M \longrightarrow B/A \longrightarrow 0$$

are in bijection with the elements of $\mathrm{Ext}^1_R(B/A, A)$; this is a topic we previously skipped.

For a concrete example, take the case where $A = \mathbb{Z}/2$ and $B/A = \mathbb{Z}/2$. Then B is not uniquely determined: $\mathrm{Ext}^1_R(\mathbb{Z}/2, \mathbb{Z}/2) \cong \mathbb{Z}/2$, and B can be either $\mathbb{Z}/4$ or $\mathbb{Z}/2 \oplus \mathbb{Z}/2$, which are not isomorphic.

8.4 Convergence of spectral sequences

But back to spectral sequences. Note that any general statement we make about cohomological spectral sequences can easily be translated into a statement about homological spectral sequences, so we will alternate between the two.

Definition 8.30. We say that a spectral sequence $E = (E^r)$ is **bounded** if for every n there are only finitely many nonzero terms of total degree n, meaning that there are only finitely many pairs (p, q) with $n = p + q$ such that $E^r_{p,q} \neq 0$ for all r.

Remark 8.31. Notice that if a spectral sequence is bounded, then for every (p, q) there exists an r such that all the differentials in and out of $E^r_{p,q}$ are zero, and thus $E^r_{p,q} = E^{r+1}_{p,q}$.

Most spectral sequences one deals with end up being bounded, so we will focus only on the case of bounded spectral sequences. This guarantees that what we are about to do makes sense.

From the data in a spectral sequence E we define a limiting page E_∞, which can often be identified with some interesting object (for example, the homology of a complex we care about). These E_∞ pages are sort of the whole point of the spectral sequence business, as they contain (pieces of) the information we want to compute.

Construction 8.32 (E_∞ page)**.** Consider a spectral sequence (E_r). Set

$$B_0 := 0 \quad \text{and} \quad Z_0 = E_0,$$

so that $E_0 = Z_0/B_0$. At each stage r, given $B_r \subseteq Z_r$ such that $E_r \cong Z_r/B_r$, we define

$$Z_{r+1} := \ker\left(Z_r \longrightarrow Z_r/B_r = E_r \xrightarrow{d_r} E_r = Z_r/B_r \right)$$

and B_{r+1} such that $B_r \subseteq B_{r+1} \subseteq Z_r$ and

$$B_{r+1}/B_r := \operatorname{im}\left(Z_r \longrightarrow Z_r/B_r = E_r \xrightarrow{d_r} E_r = Z_r/B_r \right).$$

At each stage,

$$B_r \subseteq B_{r+1} \subseteq Z_{r+1} \subseteq Z_r,$$

and

$$E_{r+1} = H(E_r) = Z_{r+1}/B_{r+1}.$$

Thus we get chains

$$0 = B_0 \subseteq B_1 \subseteq B_2 \subseteq \cdots \subseteq Z_r \subseteq Z_{r-1} \subseteq \cdots \subseteq Z_1 \subseteq Z_0 = E_0.$$

We say that Z_r consists of the elements that **survive until stage** r, while B_r consists of the elements that are in the image of the differentials by stage r. We define

$$B_\infty := \bigcup_i B_i \qquad \text{and} \qquad Z_\infty := \bigcap_i Z_i.$$

The elements in $B^{p,q}_\infty$ are the (classes of) those elements in $E^{p,q}_0$ that are in the image of the differential at some stage; we say these are the elements that are **eventually bound**. The elements in $Z^{p,q}_\infty$ are the (classes of) those elements in $E^{p,q}_0$ that are in the kernel of all the differentials at all stages, so they **survive forever** or **live forever**.

Definition 8.33. Given a homological spectral sequence $E = (E_r)$, the E_∞ page is the bigraded module given by

$$E_\infty^{p,q} := Z_\infty^{p,q}/B_\infty^{p,q}.$$

Similarly, given a cohomological spectral sequence (E^r) one can define Z^r and B^r, Z^∞ and B^∞, and the E^∞ page

$$E_{p,q}^\infty := Z_{p,q}^\infty/B_{p,q}^\infty.$$

Lemma 8.34. *Let $E = (E_r)$ be a spectral sequence. We have*

$$E_{r+1} = E_r \iff Z_{r+1} = Z_r \text{ and } B_{r+1} = B_r.$$

Proof. In general, if X/Y is a subquotient of Z, we have $Y \subseteq X \subseteq Z$, so $X/Y = Z$ if and only if $Y = 0$ and $X = Z$. If $E_{r+1} = E_r$, then

$$Z_{r+1}/B_{r+1} = E_{r+1} = E_r = Z_r/B_r,$$

so $B_{r+1} = 0$ in Z_r/B_r, so we must have $B_{r+1} = B_r$. But then

$$Z_{r+1}/B_r = Z_r/B_r,$$

so $Z_{r+1} = Z_r$.

Conversely, if $Z_{r+1} = Z_r$ and $B_{r+1} = B_r$, then

$$E_{r+1} = Z_{r+1}/B_{r+1} = Z_r/B_r = E_r. \qquad \square$$

This E^∞ is easier to compute in the following special cases.

Definition 8.35. We say that a spectral sequence $E = (E_r)$ **degenerates** at the nth page if $d_r = 0$ for all $r \geqslant n$.

Remark 8.36. If a spectral sequence degenerates at the rth page, then E_r is a limit term for the spectral sequence, in the sense that $E_s = E_r$ for all $s \geqslant r$, and $E_\infty = E_r$.

Example 8.37. Let E be a spectral sequence. If the rth page is concentrated in one row or one columnn, we say that the spectral sequence **collapses** at the rth page. Notice that in such situations the spectral sequence will automatically degenerate at the rth page.

Definition 8.38. Le H be a graded R-module. We say that a spectral sequence $E = (E^r)$ **converges** to the graded R-module $H = (H_n)$, and denote it by

$$E_{p,q}^2 \implies H_{p+q},$$

if there exists a bounded filtration $F^\bullet$ for H such that

$$E_{p,q}^\infty \cong F^p H_{p+q}/F^{p-1}H_{p+q} \quad \text{for all } p, q.$$

Alternatively, we may write $E_{p,q}^r \implies H_{p+q}$ for some other fixed choice of r besides $r = 2$.

The idea is that the spectral sequence E *approximates* the graded module H. The E^∞ page does not quite compute H exactly, but it gives us an approximation: the factors of a filtration. As we noted in Theorem 8.29, a filtration of a module does not always compute the module exactly, but it gives us some information.

Remark 8.39. Given a spectral sequence E converging to a graded module H, if the E^∞ page has only one nonzero term on the $p+q=n$ diagonal, say $E^\infty_{p,q} = A$, then $H_n = A$. Indeed, our filtration must be of the form $F^pH = 0$ for $p < 0$ and $F^pH = A$ for $p \geqslant 0$, and since we assumed that the filtration is bounded, the only option is for $H_n = A$.

For the same reason, if there are no nonzero terms on the $p+q=n$ diagonal, then we must have $H_n = 0$.

More generally, suppose that the nonzero terms on the $p+q=n$ diagonal are $A_1, \dots, A_s$, with $A_i = E^\infty_{p_i, n-p_i}$ and $p_1 < p_2 < \cdots < p_s$. The first term is the first nonzero factor in the filtration, which means that H_n must have a submodule isomorphic to A_1. The second term A_2 corresponds to a factor B_2/A_1, where B_2 is a submodule of H_n. So B_2 fits into a short exact sequence

$$0 \longrightarrow A_1 \longrightarrow B_2 \longrightarrow A_2 \longrightarrow 0.$$

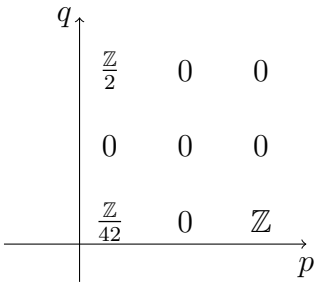
If $\operatorname{Ext}^1_R(A_2, A_1) = 0$, then there are no choices, and B_2 must be $A_1 \oplus A_2$. But in general, we might get multiple possibilities for B_2. Next we get a submodule B_3 such that $B_3/B_2 \cong A_3$, so we have a short exact sequence

$$0 \longrightarrow B_2 \longrightarrow B_3 \longrightarrow A_2 \longrightarrow 0.$$

And so on. Finally, H_n fits into a short exact sequence

$$0 \longrightarrow B_{s-1} \longrightarrow H_n \longrightarrow A_s \longrightarrow 0.$$

Example 8.40. Suppose that E is a spectral sequence such that $E^2_{p,q} \Rightarrow H_{p+1}$, with E^∞ page



The diagonal $p+q=0$ tells us that H has a filtration $F^\bullet H$ with

$$F^0H_0/F^{-1}H_0 = \mathbb{Z}/42 \quad \text{and} \quad F^nH_0/F^{n-1}H_0 = 0 \text{ for all } n \neq 0.$$

Since we assume that the filtration is bounded, this means that we our filtration is $F^pH = 0$ for $p < 0$ and $F^pH = H$ for $p \geqslant 0$, so in particular $H_0 = \mathbb{Z}/42$. In fact, whenever we have a unique nonzero term on the $p+q=n$ diagonal, that term is H_n.

Moreover, by the same reasoning we conclude that $H_n = 0$ for $n < 0$, $n > 2$ and $n = 1$, since there are no nonzero terms on the $p + q = n$ diagonals for all those values of n.

The diagonal $p + q = 2$ is the most interesting. It tells us that H_2 has a filtration

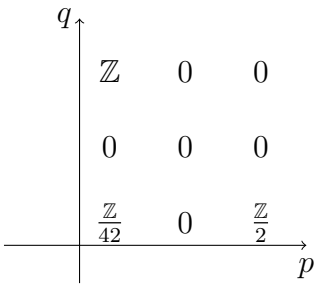
$$F^0 H_2/F^{-1} H_2 = \mathbb{Z}/2 \quad \text{and} \quad F^2 H_2/F^1 H_2 = \mathbb{Z}.$$

Since the first nonzero factor of the filtration is $\mathbb{Z}/2$, this says that H_2 has a submodule $A \cong \mathbb{Z}/2$. The next nonzero factor is also the last nonzero factor, so it tells us that $H_2/A \cong \mathbb{Z}$. Thus H_2 fits into a short exact sequence

$$0 \longrightarrow \mathbb{Z}/2 \longrightarrow H_2 \longrightarrow \mathbb{Z} \longrightarrow 0.$$

Luckily, $\mathrm{Ext}^1_{\mathbb{Z}}(\mathbb{Z}, \mathbb{Z}/2) = 0$, so the only such short exact sequence is the trivial one. We conclude that $H_2 \cong \mathbb{Z} \oplus \mathbb{Z}/2$.

Note, however, that if instead we had E^∞ page



then we could only say that H_2 has a submodule isomorphic to $\mathbb{Z}$, and that it fits into a short exact sequence of the form

$$0 \longrightarrow \mathbb{Z} \longrightarrow H_2 \longrightarrow \mathbb{Z}/2 \longrightarrow 0.$$

In particular, we have two options for H_2: we may have

$$H_2 \cong \mathbb{Z} \oplus \mathbb{Z}/2 \text{ or } H_2 \cong \mathbb{Z},$$

since

$$0 \longrightarrow \mathbb{Z} \xrightarrow{2} \mathbb{Z} \longrightarrow \mathbb{Z}/2 \longrightarrow 0.$$

is also a short exact sequence (and it does not split). We know these are all the possibilities since $\mathrm{Ext}^1_{\mathbb{Z}}(\mathbb{Z}/2, \mathbb{Z}) \cong \mathbb{Z}/2$, which has two elements.

We have the basic definitions but we haven't yet seen any examples. This is by design; to give a good example of a spectral sequence we need to do a bit more work than to give a starting example for a run of the mill definition. In the next few sections we will discuss some of the ways in which spectral sequences arise. There are, however, many interesting spectral sequences one would discuss in a first course that we do not have the time to cover. We strongly encourage the reader to seek out better sources, such as those we listed in the beginning of the chapter.

8.5 The spectral sequence of a filtered complex

We will now describe how every filtered complex gives rise to a spectral sequence. We will be writing a homological spectral sequence, but by setting slightly different conventions one could also write a cohomological spectral sequence; this part is up to your personal taste. Given a filtered complex (C, ∂, F), we first, define

$$E^0_{p,q} := F^pC_{p+q}/F^{p-1}C_{p+q}$$

and let $d_0 \colon E^0_{p,q} \longrightarrow E^0_{p,q-1}$ be the differential induced on the associated graded $\operatorname{gr}_F C$. Each column of the E^0 page contains a complex, one of the levels of the filtration.

$$\begin{array}{ccccc}
 & \vdots & \vdots & \vdots & \\
 & \downarrow & \downarrow & \downarrow & \\
 & \frac{F^{p-1}C_{p+q+1}}{F^{p-2}C_{p+q+1}} & \frac{F^pC_{p+q+1}}{F^{p-1}C_{p+q+1}} & \frac{F^{p+1}C_{p+q+1}}{F^pC_{p+q+1}} & \\
 & \downarrow \partial & \downarrow \partial & \downarrow \partial & \\
\cdots & \frac{F^{p-1}C_{p+q}}{F^{p-2}C_{p+q}} & \frac{F^pC_{p+q}}{F^{p-1}C_{p+q}} & \frac{F^{p+1}C_{p+q}}{F^pC_{p+q}} & \cdots \\
 & \downarrow \partial & \downarrow \partial & \downarrow \partial & \\
 & \frac{F^{p-1}C_{p+q-1}}{F^{p-2}C_{p+q-1}} & \frac{F^pC_{p+q-1}}{F^{p-1}C_{p+q-1}} & \frac{F^{p+1}C_{p+q-1}}{F^pC_{p+q-1}} & \\
 & \downarrow & \downarrow & \downarrow & \\
 & \vdots & \vdots & \vdots &
\end{array}$$

Next, let

$$E^1_{p,q} = \mathrm{H}_{p+q}\left(F^pC/F^{p-1}C\right)$$

and note that

$$\mathrm{H}_{p+q}\left(F^pC/F^{p-1}C\right) = \frac{\ker\left(d^0 \colon E^0_{p,q} \to E^0_{p,q-1}\right)}{\operatorname{im}\left(E^0_{p,q+1} \to E^0_{p,q}\right)} = \mathrm{H}_{p,q}(d^0 \colon E^0_{\bullet,\bullet}).$$

Thus $E^0_{\bullet,\bullet}$ and $E^1_{\bullet,\bullet}$ are consistent with being the zeroeth and first pages of a spectral sequence. Before we move on, note that we can also rewrite the E^1 page as follows:

$$E^1_{p,q} = \frac{\{x \in F^pC_{p+q} \mid \partial(x) \in F^{p-1}C_{p+q-1}\}}{F^{p-1}C_{p+q} + \partial(F^pC_{p+q+1})}.$$

One way to interpret this is that if we take an element $x \in F^pC_{p+q}$ representing the class $[x] \in G_pC_{p+q}$ such that $\partial([x]) = 0$, then the element $\partial(x) \in F^pC_{p+q-1}$ is really in $F^{p-1}C_{p+q-1}$, so we really should move it to the column on our left (the one indexed by $p-1$).

Thus the differential on $E^1_{p,q}$ is the map $d_1 \colon E^1_{p,q} \to E^1_{p-1,q}$ defined as follows: since each class $[x] \in E^1_{p,q}$ corresponds to an element $x \in F^pC_{p+q}$ such that $\partial(x) \in F^{p-1}C_{p+q-1}$, we recast $\partial(x)$ by asking about its class in $E^1_{p-1,q} = F^{p-1}C_{p+q-1}/F^{p-2}C_{p+q-1}$. If $[x] \in \ker d_1$, that means $\partial(x) \in F^{p-2}C_{p+q-1}$, and so on. Ultimately, we will gather information about how deep into our filtration we can go to still find $\partial(x)$.

Inspired by this, we define

$$E^r_{p,q} := \frac{\{x \in F^pC_{p+q} \mid \partial(x) \in F^{p-r}C_{p+q-1}\}}{F^{p-1}C_{p+q} + \partial(F^{p+r-1}C_{p+q+1})}.$$

Now given $x \in F^pC_{p+q}$ representing a class $[x] \in E^r_{p,q}$, by definition

$$\partial(x) \in F^{p-r}C_{p+q-1},$$

so we define

$$d_r([x]) = [\partial(x)] \in E^r_{p-r,q+r-1} = \frac{\{y \in F^{p-r}C_{p+q-1} \mid \partial(y) \in F^{p-2r}C_{p+q-2}\}}{F^{p-r-1}C_{p+q-1} + \partial(F^{p-1}C_{p+q})}$$

We claim that this is well-defined, and leave the details as an exercise. Moreover, $d_r^2 = 0$ since $\partial^2 = 0$. Finally, each class in the kernel of d_r is represented by some $x \in F^pC_{p+q}$ such that $\partial(x) \in F^{p-r-1}C_{p+q-1}$, while the image of $d_r \colon E^r_{p+r,q-r+1} \to E^r_{p,q}$ consists of the image of $\partial(F^{p+r}C_{p+q+1})$. Ultimately, this gives us an isomorphism

$$E^{r+1}_{p,q} \cong \mathrm{H}_{p,q}(E^r).$$

Theorem 8.41. *Let (C, ∂) be a filtered complex with filtration $F^\bullet C$, and let*

$$E^r_{p,q} := \frac{\{x \in F^pC_{p+q} \mid \partial(x) \in F^{p-r}C_{p+q-1}\}}{F^{p-1}C_{p+q} + \partial(F^{p+r-1}C_{p+q+1})}.$$

Then ∂ induces a well-defined map

$$\begin{aligned} E^r_{p,q} &\xrightarrow{d_r} E^r_{p-r,q+r-1} \\ [x] &\longmapsto [\partial(x)] \end{aligned}$$

such that $d_r d_r = 0$, and

$$E^{r+1}_{p,q} \cong \mathrm{H}_{p,q}(E^r).$$

Thus this gives us a spectral sequence with

$$E^1_{p,q} = \mathrm{H}_{p+q}(G_pC_\bullet) = \mathrm{H}_{p+1}\left(F^pC_\bullet/F^{p-1}C_\bullet\right).$$

Moreover, if the filtration is bounded, then the spectral sequence converges to

$$E^\infty_{p,q} = F^p \mathrm{H}_{p+q}(C_\bullet)/F^{p-1} \mathrm{H}_{p+q}(C_\bullet).$$

Thus

$$E^\infty_{p,q} \Rightarrow \mathrm{H}_{p+q}(C).$$

This provides a comparison between taking the homology of the associated graded module, or taking the associated graded module of the homology.

Example 8.42. Consider any short exact sequence of complexes, which we can always write as the inclusion of a subcomplex A into a complex B followed by the quotient map:

$$0 \longrightarrow A \xrightarrow{\ i\ } B \xrightarrow{\ \pi\ } A/B \longrightarrow 0.$$

It turns out that we can recover the long exact sequence in homology given by the Snake Lemma via a spectral sequence.

To do that, we can think of the inclusion $A \subseteq B$ as a filtration on $C = B$: we set

$$F^n C = \begin{cases} 0 & \text{if } n < 0 \\ A & \text{if } n = 0 \\ B & \text{if } n \geqslant 1. \end{cases}$$

Let us compute the spectral sequence of this filtered complex. First, note that In E^0 page, we have

$$E^0_{0,q} = F^0 C_q / F^{-1} C_q = A_q \quad \text{and} \quad E^0_{1,q} = F^1 C_{q+1} / F^0 C_{q+1} = B_{q+1}/A_{q+1}.$$

The differential d^0 on the E^0 page is induced by the differential ∂ on A and B. Note that we denote the differential on both complexes A and B by the same letter since the differential on A is just the restriction of the differential on B. So the E^0 page looks like

E^0

$$\begin{array}{cc} \vdots & \vdots \\ \downarrow & \downarrow \\ A_{q+1} & B_{q+2}/B_{q+2} \\ \partial \downarrow & \partial \downarrow \\ A_q & B_{q+1}/A_{q+1} \\ \partial \downarrow & \partial \downarrow \\ A_{q-1} & B_q/A_q \\ \downarrow & \downarrow \\ \vdots & \vdots \end{array}$$

where only the 0th and first columns are nonzero.

The E^1 page is obtained by taking homology, so it looks like

E^1

$$\begin{array}{c} \mathrm{H}_{q+1}(A) \xleftarrow{\ \partial\ } \mathrm{H}_{q+2}(B/A) \\ \mathrm{H}_q(A) \xleftarrow{\ \partial\ } \mathrm{H}_{q+1}(B/A) \\ \mathrm{H}_{q-1}(A) \xleftarrow{\ \partial\ } \mathrm{H}_q(B/A) \end{array}$$

and the differentials d^1, which are now horizontal, are induced by ∂. More precisely, we can rewrite the terms on the right column as

$$E^1_{1,q} = \mathrm{H}_{q+1}(B/A) = \frac{\{x \in B_{q+1} \mid \partial(x) \in A_q\}}{A_{q+1} + \partial(B_{q+2})},$$

and so by definition any class $[x] \in E^1_{q,1}$ is represented by some $x \in B_{q+1}$ such that $\partial(x) \in A_q$. Since $\partial^2 = 0$, then $\partial(x) \in Z_q(A)$, and so we can ask about the class of $\partial(x)$ in $\mathrm{H}_q(Z)$. So the differential d^1 on E^1 is given by

$$\begin{aligned} E^1_{1,q} = \tfrac{\{x \in B_{q+1} \mid \partial(x) \in A_q\}}{A_{q+1} + \partial(B_{q+2})} &\longrightarrow \mathrm{H}_q(A) \\ [x] &\longmapsto [\partial(x)]. \end{aligned}$$

Now onto the E^2 page, where

$$E^2_{1,q} = \mathrm{H}_{1,q}(E^1) = \ker\left(E^1_{1,q} \xrightarrow{d^1} E^1_{0,q} \right)$$

and

$$E^2_{0,q} = \mathrm{H}_{0,q}(E^1) = \mathrm{H}_q(A)/\operatorname{im}(d^1).$$

Note now that d^2 is a map of degree $(-2, 1)$, and thus $d^2 = 2$, since we always have zero as the source or target (or both) of d^2. Thus $E^n_{p,q} = E^2_{p,q}$ for all $n \geqslant 2$, and so by Theorem 8.41 this spectral sequence converges, with

$$E^\infty_{p,q} \cong F^p \mathrm{H}_{p+q}(C)/F^{p-1} \mathrm{H}_{p+q}(C).$$

Now $F^1 \mathrm{H}_n(C) = \mathrm{H}_n(B)$, while

$$F^0 \mathrm{H}_n(C) = \operatorname{im}\left(\mathrm{H}_n(A) \to \mathrm{H}_n(B)\right).$$

Here the map $\mathrm{H}(A) \to \mathrm{H}(B)$ is the map induced by the original inclusion of A into B. Moreover, $F^n \mathrm{H}(C) = 0$ for all $n < 0$ and $F^n \mathrm{H}(C) = \mathrm{H}(B)$ for all $n \geqslant 1$. So for all n we get isomorphisms

$$\ker\left(E^1_{1,n-1} \xrightarrow{d^1} E^1_{0,n-1} \right) \cong \mathrm{H}_n(B)/\operatorname{im}\left(\mathrm{H}_n(A) \to \mathrm{H}_n(B)\right) = \mathrm{H}_n(B)/\operatorname{im} \mathrm{H}_n(i)$$

and

$$\mathrm{H}_n(A)/\operatorname{im}(d^1) \cong \operatorname{im}\left(\mathrm{H}_n(A) \to \mathrm{H}_n(B)\right) = \operatorname{im}(\mathrm{H}_n(i)).$$

So we now have all the tools we need to construct the long exact sequence in homology from this spectral sequence. First, we construct a complex. Our map d^1 from the spectral sequence gives us homomorphisms

$$\mathrm{H}_{n+1}(B/A) \xrightarrow{d^1} \mathrm{H}_n(A).$$

By definition, $d^1([x]) = [\partial(x)]$; composing this with the map $\mathrm{H}_n(i) \colon \mathrm{H}_n(A) \to \mathrm{H}(B)$ consists of viewing $\partial(x) \in Z_n(B/A)$ and asking for its class in $\mathrm{H}_n(B)$. But $\partial(x) \in A_n \subseteq Z_n(B/A)$, so the composition

$$\mathrm{H}_{n+1}(B/A) \xrightarrow{d^1} \mathrm{H}_n(A) \xrightarrow{\mathrm{H}_n(i)} \mathrm{H}_n(B)$$

is zero.

Similarly, the map $\mathrm{H}_{n+1}(B) \to \mathrm{H}_{n+1}(A)$ takes a class $[y] \in \mathrm{H}_n(B)$, represented by $y \in Z_n(B)$, and identifies y with its image in $Z_n(B) \subseteq Z_n(B/A)$. But $\partial(y) = 0$ by definition, so $[\partial(y)] = 0$ and the composition

$$\begin{array}{ccccc} \mathrm{H}_{n+1}(B) & \longrightarrow & \mathrm{H}_{n+1}(B/A) & \xrightarrow{d^1} & \mathrm{H}_n(A) \\ [y] & \longmapsto & [y] & \longmapsto & [\partial(y)] \end{array}$$

is the zero map. Moreover, H_n is an additive functor, and thus it takes $A \to B \to B/A$ to a complex, so putting it all together we get a complex

$$\mathrm{H}_{n+1}(B/A) \xrightarrow{d^1} \mathrm{H}_n(A) \xrightarrow{\mathrm{H}_n(i)} \mathrm{H}_n(B) \xrightarrow{\mathrm{H}_n(\pi)} \mathrm{H}_n(B/A) \xrightarrow{d^1} \mathrm{H}_{n-1}(A).$$

All that is left is for us to check that exactness follows from our spectral sequence. On the one hand, the fact that this is a complex says that

$$\operatorname{im}(d^1) \subseteq \ker(\mathrm{H}_n(i)),$$

but our spectral sequence gave us the isomorphism

$$\mathrm{H}_n(A)/\operatorname{im}(d^1) \cong \operatorname{im}(\mathrm{H}_n(i)),$$

so by the First Isomorphism Theorem we get

$$\mathrm{H}_n(A)/\ker(\mathrm{H}_n(i)) \cong \operatorname{im}(\mathrm{H}_n(i)) \cong \mathrm{H}_n(A)/\operatorname{im}(d^1).$$

We can now conclude that $\operatorname{im}(d^1) = \ker \mathrm{H}_n(i)$. This gives us exactness at $\mathrm{H}_n(A)$. Similarly, we claim that the fact that we have a complex gives us a map

$$\begin{array}{ccc} \operatorname{coker}(\mathrm{H}_n(i)) & \xrightarrow{\psi} & \ker(d^1) \\ x + \operatorname{im}(\mathrm{H}_n(i)) & \longmapsto & \mathrm{H}_n(\pi)(x). \end{array}$$

Indeed, for any $x \in \mathrm{H}_n(B)$ we get $\mathrm{H}_n(\pi)(x) \in \ker d^1$, and the map is well-defined since if $x \in \operatorname{im}(\mathrm{H}_n(i)) \subseteq \ker(\mathrm{H}_n(\pi))$ then by definition $\mathrm{H}_n(\pi)(x) = 0$, so that we get a well-defined map from $\operatorname{coker}(\mathrm{H}_n(i))$. But our spectral sequence gave us an isomorphism

$$\ker(d^1) \cong \mathrm{H}_n(B)/\operatorname{im}(\mathrm{H}_n(i)) = \operatorname{coker}(\mathrm{H}_n(i)),$$

so that the map ψ we defined above is an isomorphism, and thus we must have

$$\ker(\mathrm{H}_n(\pi)) = \operatorname{im}(\mathrm{H}_n(i)),$$

proving exactness at $\mathrm{H}_n(B)$. Finally, this also says that we can rewrite our previous isomorphism as

$$\ker(d^1) \cong \mathrm{H}_n(B)/\operatorname{im}(\mathrm{H}_n(i)) = \mathrm{H}_n(B)/\ker(\mathrm{H}_n(\pi)),$$

but by the First Isomorphism Theorem we get

$$\ker(d^1) \cong \operatorname{im}(\mathrm{H}_n(\pi)).$$

We already knew that $\operatorname{im}(\mathrm{H}_n(\pi)) \subseteq \ker(d^1)$ from the fact that we had a complex, so we conclude that we must have exactness at $\mathrm{H}_n(B/A)$.

Thus we have recovered the connecting homomorphism from the Snake Lemma and the long exact sequence in homology, all via this spectral sequence.

8.6 The spectral sequence of a double complex

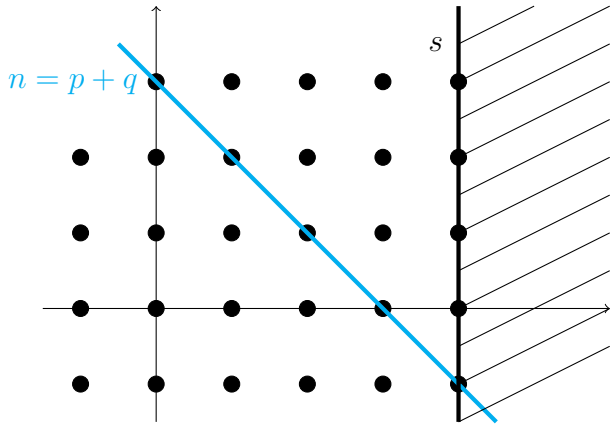
One important example of a spectral sequence induced by a filtered complex is the case of the spectral sequence of a double complex.

Definition 8.43. Let C be a double complex. There are two canonical filtrations on the total complex of C:

(I) The **First Filtration** of $\mathrm{Tot}^{\oplus}(C)$ is the filtration ${}^{I}F^{s}\,\mathrm{Tot}^{\oplus}(C)$ given by

$$ {}^{I}F^{s}\,\mathrm{Tot}^{\oplus}(C)_n := \bigoplus_{i \leqslant s} C_{i,n-i}. $$

For each s, ${}^{I}F^{s}\,\mathrm{Tot}^{\oplus}(C)$ is the subcomplex of $\mathrm{Tot}^{\oplus}(C)$ obtained by first truncating C at $p = s$:

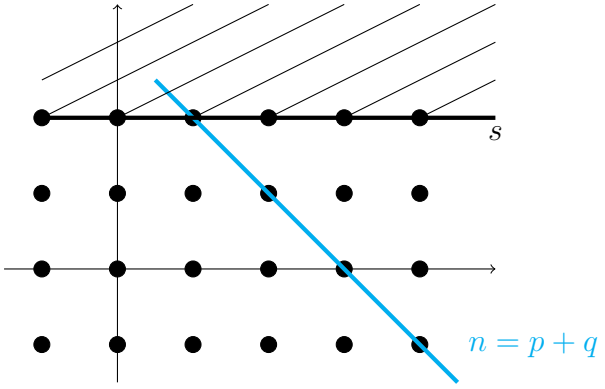


The First Filtration

(II) The **Second Filtration** of $\mathrm{Tot}^{\oplus}(C)$ is the filtration ${}^{II}F^{s}\,\mathrm{Tot}^{\oplus}(C)$ given by

$$ {}^{II}F^{s}\,\mathrm{Tot}^{\oplus}(C)_n := \bigoplus_{j \leqslant s} C_{n-j,j}. $$

For each s, ${}^{II}F^{s}\,\mathrm{Tot}^{\oplus}(C)$ is the subcomplex of $\mathrm{Tot}^{\oplus}(C)$ obtained by truncating C at $q = s$:



The Second Filtration

Each of these filtered complexes gives rise to a spectral sequence.

Discussion 8.44 (First Spectral Sequence)**.** Let us examine the spectral sequence induced by the First Filtration, which we write as ${}^{I}E^{s}$. To make the notation less heavy, we will write

$$E^r_{p,q} := {}^{I}E^r_{p,q} \quad \text{and } F^p := {}^{I}F^s \operatorname{Tot}^{\oplus}(C) \quad \text{so} \quad F^p_n = \bigoplus_{i \leqslant s} C_{i,n-i}.$$

Our spectral sequence starts with

$$E^0_{p,q} = F^p_{p+q}/F^{p-1}_{p+q} = \bigoplus_{i \leqslant p} C_{i,p+q-i} \Big/ \bigoplus_{i \leqslant p-1} C_{i,p+q-i} = C_{p,q},$$

so that

$$E^0_{p,\bullet} = C_{p,\bullet} = p\text{th column of } C.$$

Moreover,

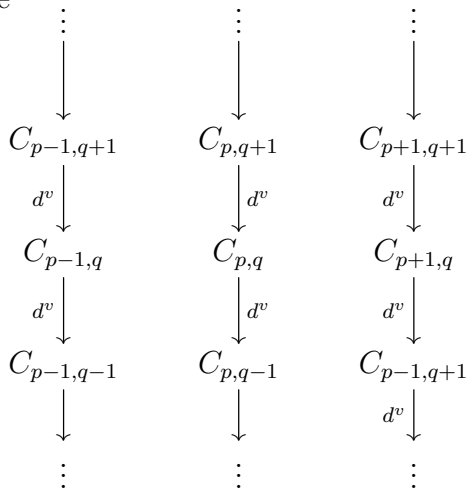
$$d_0 \colon C_{p,q} = E_0^{p,q} \longrightarrow E_0^{p,q-1} = C_{p,q-1}$$

is the map induced by the differential $d = d^v + d^h$ on this filtration. Since

$$d \colon C_{p,q} \xrightarrow{(d^h, d^v)} C_{p-1,q} \oplus C_{p,q-1},$$

we conclude that d_0 is precisely the vertical differential d^v of the original double complex C. Thus the E_0 page looks like

E^0



Then $E^1_{p,q} = \mathrm{H}^v_q(C_{p,\bullet})$, and

$$d_1 \colon \mathrm{H}^v_q(C_{p,\bullet}) = E^1_{p,q} \longrightarrow E^1_{p-1,q} = \mathrm{H}^v_q(C_{p-1,\bullet})$$

is the map induced by $d \colon C_{p,q} \xrightarrow{(d^h, d^v)} C_{p-1,q} \oplus C_{p,q-1}$. This is the horizontal differential d^h. We conclude that

$${}^{I}E^2_{p,q} = \mathrm{H}^h_p \, \mathrm{H}^v_q(C).$$

We automatically get a convergence theorem from Theorem 8.41.

Theorem 8.45. *If C is a bounded double complex, then*

$${}^{I}E^2_{p,q} \Rightarrow \mathrm{H}_{p+q}(\operatorname{Tot}^{\oplus}(C)).$$

Discussion 8.46 (Second Spectral Sequence)**.** Let us examine the spectral sequence induced by the Second Filtration, which we write as ${}^{II}E^s$. To make the notation less heavy, we will write

$$F^p := {}^{II}F^s \operatorname{Tot}^{\oplus}(C) \quad \text{so} \quad F^p_n = \bigoplus_{j \leqslant s} C_{n-j,j}.$$

Our spectral sequence starts with

$$E^0_{p,q} = F^p_{p+q}/F^{p-1}_{p+q} = \bigoplus_{j \leqslant p} C_{p+q-j,j} \Big/ \bigoplus_{j \leqslant p-1} C_{p+q-j,j} = C_{q,p},$$

so that

$$E^0_{p,\bullet} = C_{\bullet,p} = p\text{th column of } C.$$

Moreover,

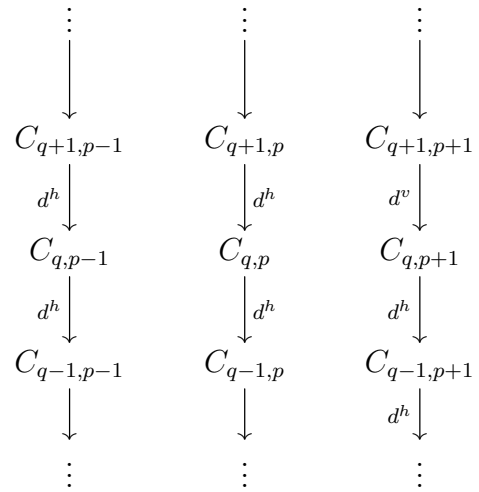
$$d_0 \colon C_{q,p} = E^0_{p,q} \longrightarrow E^0_{p,q-1} = C_{q-1,p}$$

is the map induced by the differential $d = d^v + d^h$ on this filtration. Since

$$d \colon C_{p,q} \xrightarrow{(d^h, d^v)} C_{p-1,q} \oplus C_{p,q-1} \ ,$$

we conclude that d_0 is precisely the horizontal differential d^h of the original double complex C. Thus the E_0 page looks like

E^0



Then $E^1_{p,q} = \mathrm{H}^h_q(C_{\bullet,p})$, and

$$d_1 \colon \mathrm{H}^h_q(C_{\bullet,p}) = E^1_{p,q} \longrightarrow E^1_{p-1,q} = \mathrm{H}^h_q(C_{\bullet,p-1})$$

is the map induced by $d \colon C_{p,q} \xrightarrow{(d^h, d^v)} C_{p-1,q} \oplus C_{p,q-1}$. Thus d_1 is the vertical differential d^v. We conclude that

$${}^{II}E^2_{p,q} = \mathrm{H}^v_p \, \mathrm{H}^h_q(C).$$

We now get a convergence theorem automatically from Theorem 8.41.

Theorem 8.47. *If C is a bounded double complex, then*

$${}^{II}E^2_{p,q} \Rightarrow \mathrm{H}_{p+q}(\operatorname{Tot}^{\oplus}(C)).$$

Remark 8.48. We know exactly how each spectral sequence converges to $\mathrm{H}_n(\mathrm{Tot}^{\oplus}(C))$: via the filtration on $\mathrm{H}_n(\mathrm{Tot}^{\oplus}(C))$ induced by the given filtration on $\mathrm{Tot}^{\oplus}(C)$. More precisely,

$$ {}^{I}E^{\infty}_{p,q} = \frac{{}^{I}F^{p}\,\mathrm{H}_{p+q}(\mathrm{Tot}^{\oplus}(C))}{{}^{I}F^{p-1}\,\mathrm{H}_{p+q}(\mathrm{Tot}^{\oplus}(C))} \quad \text{and} \quad {}^{II}E^{\infty}_{p,q} = \frac{{}^{II}F^{p}\,\mathrm{H}_{p+q}(\mathrm{Tot}^{\oplus}(C))}{{}^{II}F^{p-1}\,\mathrm{H}_{p+q}(\mathrm{Tot}^{\oplus}(C))}. $$

Remark 8.49. If C is a first quadrant double complex, so that $C_{i,j} = 0$ whenever $i < 0$ or $j < 0$, then for all $p < 0$ we have

$$ {}^{I}F^{p}\,\mathrm{Tot}^{\oplus}(C) = 0 \text{ and } {}^{II}F^{p}\,\mathrm{Tot}^{\oplus}(C) = 0. $$

Thus

$$ {}^{I}F^{p}\,\mathrm{H}(\mathrm{Tot}^{\oplus}(C)) = 0 \text{ and } {}^{II}F^{p}\,\mathrm{H}(\mathrm{Tot}^{\oplus}(C)) = 0. $$

Moreover, for each fixed n if we take $p > n$ or $q > n$ then

$$ {}^{I}F^{p}\,\mathrm{Tot}^{\oplus}(C)_n = \mathrm{Tot}^{\oplus}(C)_n \text{ and } {}^{II}F^{p}\,\mathrm{Tot}^{\oplus}(C)_n = \mathrm{Tot}^{\oplus}(C)_n, $$

so

$$ {}^{I}F^{p}\,\mathrm{H}_n(\mathrm{Tot}^{\oplus}(C)) = \mathrm{H}_n(\mathrm{Tot}^{\oplus}(C)) \text{ and } {}^{II}F^{p}\,\mathrm{H}_n(\mathrm{Tot}^{\oplus}(C)) = \mathrm{H}_n(\mathrm{Tot}^{\oplus}(C)). $$

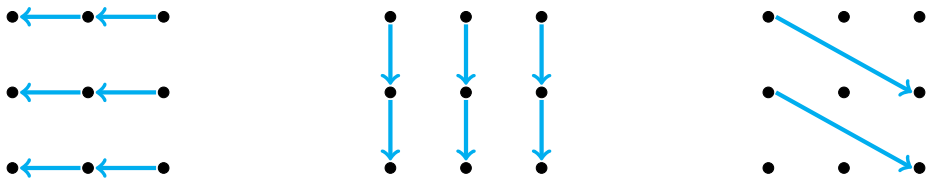
Thus each fixed diagonal of ${}^{I}E^{p,q}_{\infty}$ and ${}^{II}E^{p,q}_{2}$ contains the factors of a finite filtration on $\mathrm{H}_n(\mathrm{Tot}^{\oplus}(C))$:

$$ 0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_s = \mathrm{H}_n(\mathrm{Tot}^{\oplus}(C)). $$

In the special case when ${}^{I}E$ or ${}^{II}E$ collapses, meaning that E^{∞} is concentrated in one row or column, then the filtration on $\mathrm{H}(\mathrm{Tot}^{\oplus}(C))$ has a unique term in each degree, so that we can read $\mathrm{H}(\mathrm{Tot}^{\oplus}(C))$ exactly from the E^{∞} page.

Remark 8.50. The fact that ${}^{II}E^{0}_{p,q} = C_{q,p}$ can lead to a lot of confusion. To make things easier, one often breaks the rules a little and instead takes the E_0 page to be ${}^{II}E^{0}_{p,q} = C_{p,q}$ with *horizontal* differentials, the E^1 page to have *vertical* differentials induced by the vertical differentials on C, and then the E^r page to have a differential of degree $(r-1, -r)$. This gives us a gadget that isn't quite a spectral sequence under our formal definition, but *should* be a spectral sequence. In fact, this new convention makes everything *much* easier to read.

Under this reasonable convention, the differentials look as follows:



0th page 1st page 2nd page

Also, note that now

$$ {}^{II}E^{2}_{p,q} = \mathrm{H}^{v}_{q}\,\mathrm{H}^{h}_{p}(C), $$

and that while

$$ {}^{II}E^{\infty}_{p,q} \implies \mathrm{H}_{p+q}(\mathrm{Tot}^{\oplus}(C)), $$

we now should carefully read the filtration backwards:

$$ {}^{II}E^{\infty}_{p,q} = \frac{F^{q}\,\mathrm{H}_{p+q}(\mathrm{Tot}^{\oplus}(C))}{F^{q-1}\,\mathrm{H}_{p+q}(\mathrm{Tot}^{\oplus}(C))}. $$

Example 8.51. Let us give a new proof of the Snake Lemma using spectral sequences. More precisely, we will show that given any commutative diagram with exact rows

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{i} & B & \xrightarrow{\pi} & C & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle h} & & \\ 0 & \longrightarrow & A' & \xrightarrow[i']{} & B' & \xrightarrow[\pi']{} & C' & \longrightarrow & 0. \end{array}$$

then there exists an exact sequence

$$0 \longrightarrow \ker f \longrightarrow \ker g \longrightarrow \ker h \xrightarrow{\partial} \operatorname{coker} f \longrightarrow \operatorname{coker} g \longrightarrow \operatorname{coker} h \longrightarrow 0.$$

To do that, we start by viewing our commutative diagram as a double complex M, as follows: we set $M_{0,0} = C$, $M_{0,1} = C'$, and so on, resulting in

$$\begin{array}{ccccc} C' & \xleftarrow{\pi'} & B' & \xleftarrow{i'} & C' \\ \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle f} \\ C & \xleftarrow{\pi} & B & \xleftarrow{i} & C. \end{array}$$

Now notice that M is a first quadrant double complex with exact rows, so by the Acyclic Assembly Lemma, $\operatorname{Tot}^{\oplus}(M)$ is exact. We can also prove that $\operatorname{Tot}^{\oplus}(M)$ is exact by computing the spectral sequence arising from the Second Filtration on M, which under our new and improved notation from Theorem 8.50 has

$$E^0 = \begin{array}{ccccc} & & C & \xleftarrow{\pi} & B & \xleftarrow{i} & C \\ & & & & & & \\ & & C' & \xleftarrow[\pi']{} & B' & \xleftarrow[i']{} & C' \end{array}$$

and since the rows are all exact, we see that in the next step we will end up with $E^1 = 0$, and thus $E^\infty = 0$. Therefore, $\mathrm{H}(\operatorname{Tot}^{\oplus}(M)) = 0$.

Now consider the spectral sequence induced by the First Filtration on M, which has

$$E^0 = \begin{array}{ccc} C' & B' & C' \\ \downarrow{\scriptstyle h} & \downarrow{\scriptstyle g} & \downarrow{\scriptstyle f} \\ C & B & C \end{array} \qquad E^1 = \begin{array}{ccccc} \ker h & \xleftarrow{\pi'} & \ker g & \xleftarrow{i'} & \ker f \\ & & & & \\ \operatorname{coker} h & \xleftarrow{\pi} & \operatorname{coker} g & \xleftarrow{i} & \operatorname{coker} f. \end{array}$$

Now whatever E^2 is, since d^2 is a $(-2, 1)$ degree map, the only possible nonzero differential on E^2 is $d^2 : E^2_{2,0} \to E^2_{0,1}$. Moreover, all the differentials on all the higher pages vanish, simply because there are not enough nonzero modules already, so $E^3 = E^\infty$. But we know that $\operatorname{Tot}^{\oplus}(M)$ is exact, so in fact since our filtrations are finite we must necessarily have $E^\infty = 0$. In particular, the only potentially nonzero objects in E^2 are $E^2_{0,1}$ and $E^2_{2,0}$, so this proves the exactness of

$$\ker h \xleftarrow{\pi'} \ker g \xleftarrow{i'} \ker f \longleftarrow 0 \quad \text{and} \quad 0 \longleftarrow \operatorname{coker} h \xleftarrow{\pi} \operatorname{coker} g \xleftarrow{i} \operatorname{coker} f.$$

Moreover, the fact that $E^2_{0,1}$ and $E^2_{2,0}$ are the only possible nonzero objects in E^2 together with the fact that and $E^3_{0,1} = 0 = E^3_{2,0}$ imply that $d^2 \colon E^2_{0,1} \to E^2_{2,10}$ must be an isomorphism. More precisely, we get an isomorphism

$$\operatorname{coker}\left(\ \ker h \xleftarrow{\pi'} \ker g\ \right) = E^2_{0,1} \cong E^2_{2,0} = \ker\left(\ \operatorname{coker} g \xleftarrow{i} \operatorname{coker} f\ \right).$$

Thus we get an exact sequence

$$\begin{array}{ccccccccc} \operatorname{coker} g & \xleftarrow{i} & \operatorname{coker} f & \leftarrow & \cdots & \xleftarrow{\partial} & \cdots & \xleftarrow{\pi'} & \ker g \\ & & \nwarrow & & & & \nearrow & & \\ & & & E^2_{2,0} & \xleftarrow[\cong]{} & E^2_{0,1} & & & \end{array}$$

where the map ∂ obtained by composition is the connecting homomorphism we dreamed of.

Exercises

Exercise 8.52. Show that if M is a first quadrant double complex with exact rows or exact columns, then $\mathrm{Tot}^{\oplus}(M)$ is exact.

Exercise 8.53. Let M and N be R-modules and fix projective resolutions P for M and Q for N. Consider the double complex $C = P \otimes_R Q$.

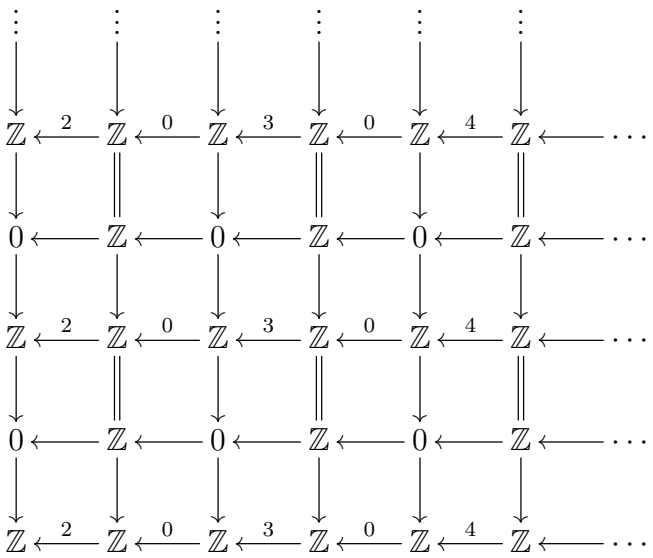
a) Compute the spectral sequence associated to the First Filtration (I) on C up to the E^2 page.

b) Compute the spectral sequence associated to the Second Filtration (II) on C up to the E^2 page.

c) Give a new proof that Tor is balanced:

$$\mathrm{H}_n(P \otimes_R N) \cong \mathrm{H}_n(M \otimes_R Q).$$

Exercise 8.54. Consider the following first quadrant double complex C:



a) Compute the spectral sequence associated to the First Filtration (I) until it stabilizes.

b) Compute the spectral sequence associated to the Second Filtration (II) until it stabilizes.

c) How are the two resulting E^∞ pages even possible considering they both converge to the same thing?

d) Compute $\mathrm{H}_3(\mathrm{Tot}^{\oplus}(M))$ explicitly using only the two spectral sequences you calculated.

e) Check your work by computing $\mathrm{H}_3(\mathrm{Tot}^{\oplus}(M))$ explicitly from the definition of $\mathrm{Tot}^{\oplus}(M)$.

Appendix A

Rings and modules

We will study complexes of R-modules; to make sure we are all speaking the same language, we record here our basic assumptions on rings and modules. You can learn more about the basic theory of rings and modules in any introductory algebra book, such as [DF04].

A.1 Rings and why they have 1

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 A.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 $1 \neq 0$ and $1 \cdot a = a = a \cdot 1$ 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 $a \cdot b$.

Note that the requirement that $1 \neq 0$ makes it so that the *zero ring* is not a ring.

Definition A.2. A ring R is a **commutative ring** if for all $a, b \in R$ we have $a \cdot b = b \cdot a$.

Definition A.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 A.4. A commutative ring R is a **domain**, sometimes called an **integral domain**, if it has no zerodivisors: $ab = 0 \Rightarrow a = 0$ or $b = 0$. Note that in particular we reserve the word domain for commutative rings.

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 A.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 A.6. For any ring R, there exists a unique homomorphism $\mathbb{Z} \to R$.

Definition A.7. 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 A.8. 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 A.9. 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 $I \neq R$.

A.2 Modules

You can learn more about the basic theory of (commutative) rings and R-modules in any introductory algebra book, such as [DF04].

Definition A.10. 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 A.11. 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.

The definitions of submodule, quotient of modules, and homomorphism of modules are very natural and easy to guess, but here they are.

Definition A.12. If $N \subseteq M$ are R-modules with compatible structures, we say that N is a **submodule** of M.

A map $M \xrightarrow{f} N$ between R-modules is a **homomorphism of R-modules** if it is a homomorphism of abelian groups that preserves the R-action, meaning $f(ra) = rf(a)$ for all $r \in R$ and all $a \in M$. We sometimes refer to R-module homomorphisms as **R-module maps**, or **maps of R-modules**. An isomorphism of R-modules is a bijective homomorphism, which we really should think about as a relabeling of the elements in our module. If two modules M and N are isomorphic, we write $M \cong N$.

Given an R-module M and a submodule $N \subseteq M$, the **quotient** M/N is an R-module whose elements are the equivalence classes determined by the relation on M given by $a \sim b \Leftrightarrow a - b \in N$. One can check that this set naturally inherits an R-module structure from the R-module structure on M, and it comes equipped with a natural **canonical map** $M \longrightarrow M/N$ induced by sending 1 to its equivalence class.

Example A.13. The modules over a field k are precisely all the k-vector spaces. Linear transformations are precisely all the k-module maps.

While vector spaces make for a great first example, be warned that many of the basic facts we are used to from linear algebra are often a little more subtle in commutative algebra. These differences are features, not bugs.

Example A.14. The $\mathbb{Z}$-modules are precisely all the abelian groups.

Example A.15. When we think of the ring R as a module over itself, the submodules of R are precisely the ideals of R.

Theorem A.16 (First Isomorphism Theorem)**.** *Any R-module homomorphism $M \xrightarrow{f} N$ satisfies $M/\ker f \cong \operatorname{im} f$.*

The first big noticeable difference between vector spaces and more general R-modules is that while every vector space has a basis, most R-modules do not.

Definition A.17. A subset $\Gamma \subseteq M$ of an R-module M is a **generating set**, or a **set of generators**, if every element in M can be written as a finite linear combination of elements in M with coefficients in R. A **basis** for an R-module M is a generating set Γ for M such that $\sum_i a_i\gamma_i = 0$ implies $a_i = 0$ for all i. An R-module is **free** if it has a basis.

Remark A.18. Every vector space is a free module.

Remark A.19. Every free R-module is isomorphic to a direct sum of copies of R. Indeed, let's construct such an isomorphism for a given free R-module M. Given a basis $\Gamma = \{\gamma_i\}_{i\in I}$ for M, let

$$\bigoplus_{i\in I} R \xrightarrow{\pi} M \quad .$$

$$(r_i)_{i\in I} \longrightarrow \sum_i r_i\gamma_i$$

The condition that Γ is a basis for M can be restated into the statement that π is an isomorphism of R-modules.

One of the key things that makes commutative algebra so rich and beautiful is that most modules are in fact *not* free. In general, every R-module has a generating set — for example, M itself. Given some generating set Γ for M, we can always repeat the idea above and write a **presentation** $\oplus_{i\in I} R \xrightarrow{\pi} M$ for M, but in general the resulting map π will have a nontrivial kernel. A nonzero kernel element $(r_i)_{i\in I} \in \ker\pi$ corresponds to a **relation** between the generators of M.

Remark A.20. Given a set of generators for an R-module M, any homomorphism of R-modules $M \longrightarrow N$ is determined by the images of the generators.

We say that a module is **finitely generated** if we can find a finite generating set for M. The simplest finitely generated modules are the cyclic modules.

Example A.21. An R-module is **cyclic** if it can be generated by one element. Equivalently, we can write M as a quotient of R by some ideal I. Indeed, given a generator m for M, the kernel of the map $R \xrightarrow{\pi} M$ induced by $1 \mapsto m$ is some ideal I. Since we assumed that m generates M, π is automatically surjective, and thus induces an isomorphism $R/I \cong M$.

Similarly, if an R-module has n generators, we can naturally think about it as a quotient of R^n by the submodule of relations among those n generators.

Index

Bibliography

[DF04] David S. Dummit and Richard M. Foote. *Abstract algebra*. Wiley, 3rd ed edition, 2004.

[Eis95] David Eisenbud. *Commutative algebra with a view toward algebraic geometry*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.

[Eis05] David Eisenbud. *The geometry of syzygies*, volume 229 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005. A second course in commutative algebra and algebraic geometry.

[Fre03] Peter J. Freyd. Abelian categories [mr0166240]. *Repr. Theory Appl. Categ.*, (3):1–190, 2003.

[Hat02] Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.

[ILL+07] Srikanth B. Iyengar, Graham J. Leuschke, Anton Leykin, Claudia Miller, Ezra Miller, Anurag K. Singh, and Uli Walther. *Twenty-four hours of local cohomology*, volume 87 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2007.

[Kap58] Irving Kaplansky. Projective modules. *Ann. of Math (2)*, 68:372–377, 1958.

[Lyu92] Gennady Lyubeznik. The number of defining equations of affine algebraic sets. *American Journal of Mathematics*, 114(2):413–463, 1992.

[Mac50] Saunders MacLane. Duality for groups. *Bulletin of the American Mathematical Society*, 56(6):485 – 516, 1950.

[Mat58] Eben Matlis. Injective modules over Noetherian rings. *Pacific J. Math.*, 8:511–528, 1958.

[Mat89] Hideyuki Matsumura. *Commutative ring theory*, volume 8 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 1989. Translated from the Japanese by M. Reid.

[McC01] John McCleary. *A user's guide to spectral sequences*, volume 58 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[ML98] Saunders Mac Lane. *Categories for the working mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1998.

[Pee11] Irena Peeva. *Graded syzygies*, volume 14 of *Algebra and Applications*. Springer-Verlag London, Ltd., London, 2011.

[Rie17] E. Riehl. *Category Theory in Context*. Aurora: Dover Modern Math Originals. Dover Publications, 2017.

[Rot09] Joseph J. Rotman. *An introduction to homological algebra*. Universitext. Springer, New York, second edition, 2009.

[Wei94] Charles A. Weibel. *An introduction to homological algebra*, volume 38 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1994.