

Homological Algebra

Math 915 Fall 2023

Eloísa Grifo
September 16, 2023

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 are heavily influenced by 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 Brandon Massaro, Rahul Rajkumar, Adam Richardson, Khoa Ta, Ryan Watson, and Noble Williamson, who found typos and errors. Thank you also to the students who found typos in this new version: Stephen Stern, Danny Anderson, Taylor Murray, Kesavan M, and Sara Mueller.

Contents

0	Where are we going?	1
1	Categories for the working homological algebraist	5
1.1	Categories	5
1.2	Functors	12
1.3	Natural transformations	16
1.4	The Yoneda Lemma	20
1.5	Products and coproducts	25
1.6	Limits and colimits	29
1.7	Universal properties	40
1.8	Adjoint functors	43
2	The category of chain complexes	46
2.1	Maps of complexes	46
2.2	Short exact sequences	51
2.3	Long exact sequences	57
A	Rings and modules	65
A.1	Rings and why they have 1	65
A.2	Modules	67
	Index	70

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 ∂_\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 $\text{im } \partial_{n+1} \subseteq \ker \partial_n$.

Definition 0.3. The complex $(C_\bullet, \partial_\bullet)$ is **exact** at n if $\text{im } \partial_{n+1} = \ker \partial_n$. 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 $\text{im } f_n = \ker f_{n+1}$ 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
- $\text{im } f = \ker g$.

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

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

Remark 0.7. We can rephrase Remark 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 $C = \text{coker } f$.

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 $\iota : I \rightarrow R$ and the canonical projection $\pi : R \rightarrow R/I$ 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

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

The n th **homology** of $(C_\bullet, \partial_\bullet)$ is $H_n(C)$. 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 $H^n(C)$; 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 $H_n(C_\bullet) = 0$.

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

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

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

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

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

$$\begin{aligned} H_n(C) &= 0 && \text{for } n \geq 3 \\ H_2(C) &= \frac{\ker(\mathbb{Z} \xrightarrow{4} \mathbb{Z})}{\operatorname{im}(0 \longrightarrow \mathbb{Z})} = \frac{0}{0} = 0 \\ H_1(C) &= \frac{\ker(\mathbb{Z} \xrightarrow{\pi} \mathbb{Z}/2\mathbb{Z})}{\operatorname{im}(\mathbb{Z} \xrightarrow{4} \mathbb{Z})} = \frac{2\mathbb{Z}}{4\mathbb{Z}} \cong \mathbb{Z}/2\mathbb{Z} \\ H_0(C) &= \frac{\ker(\mathbb{Z}/2\mathbb{Z} \longrightarrow 0)}{\operatorname{im}(\mathbb{Z} \longrightarrow \mathbb{Z}/2\mathbb{Z})} = \frac{\mathbb{Z}/2\mathbb{Z}}{\mathbb{Z}/2\mathbb{Z}} = 0 \\ H_n(C) &= 0 && \text{for } n < 0 \end{aligned}$$

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.

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

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

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

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

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

These ingredients satisfy the following axioms:

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

Exercise 1. Prove that every element in a category has a unique identity morphism.

Here are some categories you have likely encountered before:

Example 1.3.

- 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 1.4. 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 1.5. The category **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 $f: X \rightarrow Y$ such that $f(x) = y$, with the usual composition of functions.

Example 1.6. 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 1.7. A category \mathcal{C} is **locally small** if for all objects A and B in \mathcal{C} , $\text{Hom}_{\mathcal{C}}(A, B)$ is a set. A category \mathcal{C} is **small** if it is locally small and the collection of all objects in \mathcal{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 1.8. Given a partially ordered set (X, \leq) , we can regard X itself as a category: the objects are the elements of X , and for each x and y in X , $\text{Hom}_X(x, y)$ is either a singleton if $x \leq y$ or empty if $x \not\leq y$. There is only one possible way to define composition, and the transitive property of \leq guarantees that the composition of arrows is indeed well-defined: if there is an arrow $i \rightarrow j$ and an arrow $j \rightarrow k$, then $i \leq j$ and $j \leq k$, so $i \leq k$ and thus there is a unique arrow $i \rightarrow k$. 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 **PO**(X).

Example 1.9. For each positive integer n , the category \mathbf{n} has n objects $0, 1, \dots, n-1$ and $\text{Hom}(i, j)$ is either empty if $i > j$ or a singleton if $i \leq j$. As Example 1.8, composition is defined in the only way possible, and things work out. This is the poset category for the poset $(\{0, 1, \dots, n-1\}, \leq)$ with the usual \leq .

Example 1.10. Fix a field k . We define a category $\mathbf{Mat}\text{-}k$ with objects all positive integers, and given two positive integers a and b , the Hom-set $\text{Hom}(a, b)$ consists of all $b \times a$ matrices with entries in k . The composition rule is given by product of matrices: given $A \in \text{Hom}(a, b)$ and $B \in \text{Hom}(b, c)$, the composition $B \circ A$ is the matrix $BA \in \text{Hom}(a, c)$. For each object a , its identity arrow is given by the $a \times a$ identity matrix.

Example 1.11. 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 1.12. 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 1.13. A **diagram** in a category \mathcal{C} is a directed multigraph whose vertices are objects in \mathcal{C} and whose arrows/edges are morphisms in \mathcal{C} . A commutative diagram in \mathcal{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 1.14. The diagram

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

commutes if and only if $gf = vu$.

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

Definition 1.15. Let \mathcal{C} be any category.

- An arrow $f \in \text{Hom}_{\mathcal{C}}(A, B)$ is **left invertible** if there exists $g \in \text{Hom}_{\mathcal{C}}(B, A)$ 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

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

commutes.

²Some authors prefer the term monoid.

- An arrow $f \in \text{Hom}_{\mathcal{C}}(A, B)$ is **right invertible** if there exists $g \in \text{Hom}_{\mathcal{C}}(B, A)$ 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

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

commutes.

- An arrow $f \in \text{Hom}_{\mathcal{C}}(A, B)$ is an **isomorphism** if there exists $g \in \text{Hom}_{\mathcal{C}}(B, A)$ 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 $A \rightarrow B$.
- An arrow $f \in \text{Hom}(B, C)$ is **monic**, a **monomorphism**, or a **mono** if for all arrows

$$A \begin{array}{c} \xrightarrow{g_1} \\ \xrightarrow{g_2} \end{array} B \xrightarrow{f} C$$

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

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

$$A \xrightarrow{f} B \begin{array}{c} \xrightarrow{g_1} \\ \xrightarrow{g_2} \end{array} C$$

if $g_1f = g_2f$ then $g_1 = g_2$.

Here are some examples:

Exercise 2. Show that in **Set**, the monos coincide with the injective functions and the epis coincide with the surjective functions.

Example 1.16.

- In **Grp**, **Ring**, and **R-Mod** the isomorphisms are the morphisms that are bijective functions.
- 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 3. Show that in any category, every isomorphism is both epi and mono.

Exercise 4. Show that the usual inclusion $\mathbb{Z} \rightarrow \mathbb{Q}$ 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 5. Show that the canonical projection $\mathbb{Q} \rightarrow \mathbb{Q}/\mathbb{Z}$ is a mono in the category of divisible abelian groups.³

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

Exercise 6. 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 1.17. Let \mathcal{C} be a category. An **initial object** in \mathcal{C} is an object i such that for every object x in \mathcal{C} , $\text{Hom}_{\mathcal{C}}(i, x)$ is a singleton, meaning there exists a unique arrow $i \rightarrow x$. A **terminal object** in \mathcal{C} is an object t such that for every object x in \mathcal{C} , $\text{Hom}_{\mathcal{C}}(x, t)$ is a singleton, meaning there exists a unique arrow $x \rightarrow t$. A **zero object** in \mathcal{C} is an object that is both initial and terminal.

Exercise 7. 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 1.18.

- 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 1.19. A **subcategory** \mathcal{C} of a category \mathcal{D} consists of a subcollection of the objects of \mathcal{D} and a subcollection of the morphisms of \mathcal{D} such that the following hold:

- For every object C in \mathcal{C} , the arrow $1_C \in \text{Hom}_{\mathcal{D}}(C, C)$ is an arrow in \mathcal{C} .
- For every arrow in \mathcal{C} , its source and target in \mathcal{D} are objects in \mathcal{C} .
- For every pair of arrows f and g in \mathcal{C} such that fg is an arrow that makes sense in \mathcal{D} , fg is an arrow in \mathcal{C} .

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

Example 1.20. The category of finitely generated R -modules with R -module homomorphisms is a subcategory of **R-Mod**.

Definition 1.21. A subcategory \mathcal{C} of \mathcal{D} is a **full subcategory** if \mathcal{C} includes *all* of the arrows in \mathcal{D} between any two objects in \mathcal{C} .

Example 1.22.

- 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 1.23. Let \mathcal{C} be a category. The **opposite category** of \mathcal{C} , denoted \mathcal{C}^{op} , is a category whose objects are the objects of \mathcal{C} , and such that each arrow $f \in \text{Hom}_{\mathcal{C}^{\text{op}}}(A, B)$ is the same as some arrow in $\text{Hom}_{\mathcal{C}}(B, A)$. The composition fg of two morphisms f and g in \mathcal{C}^{op} is defined as the composition gf in \mathcal{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 \mathcal{C} and in the opposite category \mathcal{C}^{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 \mathcal{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 *cohomology*. Note that the dual of the dual is the original statement; formally, $(\mathcal{C}^{\text{op}})^{\text{op}} = \mathcal{C}$. 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) $i : \emptyset \rightarrow X$, but there are no functions $X \rightarrow \emptyset$, so $i^{\text{op}} : \emptyset \rightarrow X$ is not a function. Thus thinking about **Set**^{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 1.24. Let \mathcal{C} and \mathcal{D} be categories. A **covariant functor** $F: \mathcal{C} \rightarrow \mathcal{D}$ is a mapping that assigns to each object A in \mathcal{C} an object $F(A)$ in \mathcal{D} , and to each arrow $f \in \text{Hom}_{\mathcal{C}}(A, B)$ an arrow $F(f) \in \text{Hom}_{\mathcal{D}}(F(A), F(B))$, such that

- F preserves the composition of maps, meaning $F(fg) = F(f)F(g)$ for all arrows f and g in \mathcal{C} , and
- F preserves the identity arrows, meaning $F(1_A) = 1_{F(A)}$ for all objects A in \mathcal{C} .

A **contravariant functor** $F: \mathcal{C} \rightarrow \mathcal{D}$ is a mapping that assigns to each object A in \mathcal{C} an object $F(A)$ in \mathcal{D} , and to each arrow $f \in \text{Hom}_{\mathcal{C}}(A, B)$ an arrow $F(f) \in \text{Hom}_{\mathcal{D}}(F(B), F(A))$, such that

- F preserves the composition of maps, meaning $F(fg) = F(g)F(f)$ for all composable arrows f and g in \mathcal{C} , and
- F preserves the identity arrows, meaning $F(1_A) = 1_{F(A)}$ for all objects A in \mathcal{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 \mathcal{C} to the opposite category of \mathcal{D} , \mathcal{D}^{op} .

Remark 1.25. A contravariant functor $F: \mathcal{C} \rightarrow \mathcal{D}$ can be thought of as a covariant functor $\mathcal{C}^{\text{op}} \rightarrow \mathcal{D}$, or also as a covariant functor $\mathcal{C} \rightarrow \mathcal{D}^{\text{op}}$. 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 $F: \mathcal{C} \rightarrow \mathcal{D}$ and $G: \mathcal{D} \rightarrow \mathcal{E}$ is the same as applying a functor $\mathcal{C} \rightarrow \mathcal{D}$, which we can write as GF .

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

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

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

Exercise 8. Show that functors preserve isomorphisms.

Remark 1.26. Any functor sends isos to isos, since it preserves compositions and identities.

Example 1.27. 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 $\mathbf{Gr} \rightarrow \mathbf{Set}$ sends each group to its underlying set, and each group homomorphism to the corresponding function of sets.
- b) The identity functor $1_{\mathcal{C}}$ on any category \mathcal{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 \mathcal{C} , the **constant functor** at C is the functor $\Delta C : \mathcal{C} \rightarrow \mathcal{C}$ that sends every object to C every arrow to 1_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 $G^{\text{ab}} := G/[G, G]$ is called the **abelianization** of G . The group G^{ab} is abelian. Given a group homomorphism $f: G \rightarrow H$, f automatically takes commutators to commutators, so it induces a homomorphism $\tilde{f}: G^{\text{ab}} \rightarrow H^{\text{ab}}$. More precisely, abelianization gives a covariant functor from **Grp** to **Ab**.

- e) The unit group functor $-^*: \mathbf{Ring} \rightarrow \mathbf{Grp}$ 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 $f: R \rightarrow S$ 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 $R^* \rightarrow S^*$ 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 $\varphi: W \rightarrow V$ is a linear transformation and $\ell: V \rightarrow k$ is an element of V^* , then $\ell \circ \varphi: W \rightarrow k$ is in W^* . Doing this for all elements $\ell \in V^*$ gives a function $\varphi^*: V^* \rightarrow W^*$, 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 $\mathbf{Vect}\text{-}k \rightarrow \mathbf{Vect}\text{-}k$.
- 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 functor $R\text{-mod} \rightarrow W^{-1}R\text{-mod}$: this functor sends each R -module M to $W^{-1}M$, and each R -module homomorphism $\alpha: M \rightarrow N$ to the R -module homomorphism $W^{-1}\alpha: W^{-1}M \rightarrow W^{-1}N$.

Remark 1.28. If we apply a covariant functor to a diagram, then we get a diagram of the same shape:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ u \downarrow & & \downarrow g \\ C & \xrightarrow{v} & D \end{array} \quad \xrightarrow{\sim F} \quad \begin{array}{ccc} F(A) & \xrightarrow{F(f)} & F(B) \\ F(u) \downarrow & & \downarrow F(g) \\ F(C) & \xrightarrow{F(v)} & F(D) \end{array}$$

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

$$\begin{array}{ccc}
 A & \xrightarrow{f} & B \\
 u \downarrow & & \downarrow g \\
 C & \xrightarrow{v} & D
 \end{array}
 \quad \xrightarrow{\sim F} \quad
 \begin{array}{ccc}
 F(A) & \xleftarrow{F(f)} & F(B) \\
 F(u) \uparrow & & \uparrow F(g) \\
 F(C) & \xleftarrow{F(v)} & F(D)
 \end{array}$$

Definition 1.29. 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 1.30. A covariant functor $F : \mathcal{C} \rightarrow \mathcal{D}$ between locally small categories is

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

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

are injective.

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

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

are surjective.

- **fully faithful** if it is full and faithful.
- **essentially surjective** if every object d in \mathcal{D} is isomorphic to $F(c)$ for some c in \mathcal{C} .
- an **embedding** if it is fully faithful and injective on objects.

Example 1.31. The forgetful functor $R\text{-Mod} \rightarrow \mathbf{Set}$ 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 1.32. A fully faithful functor is not necessarily injective on objects, but it is injective on objects up to isomorphism.

Remark 1.33. A subcategory \mathcal{C} of \mathcal{D} is full if the inclusion functor $\mathcal{C} \rightarrow \mathcal{D}$ is full.

Exercise 9. Show that every fully faithful functor $F : \mathcal{C} \rightarrow \mathcal{D}$ reflects isos:

- If f is an arrow in \mathcal{C} such that $F(f)$ is an iso, then f is an iso.
- If $F(X)$ and $F(Y)$ are isomorphic, then the objects X and Y are isomorphic in \mathcal{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 1.34. Let \mathcal{C} be a locally small category. An object A in \mathcal{C} induces two Hom functors:

- The covariant functor $\text{Hom}_{\mathcal{C}}(A, -) : \mathcal{C} \rightarrow \mathbf{Set}$ is defined as follows:

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

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

$$\begin{array}{ccc}
 A & \xrightarrow{g} & B \\
 & \searrow f_*(g)=fg & \downarrow f \\
 & & C
 \end{array}$$

$f_* = \text{Hom}_{\mathcal{C}}(A, f) :$

- The contravariant functor $\text{Hom}_{\mathcal{C}}(-, B) : \mathcal{C} \rightarrow \mathbf{Set}$ is defined as follows:

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

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

$$\begin{array}{ccc}
 A & \xrightarrow{f} & C \\
 & \searrow f^*(g)=gf & \downarrow g \\
 & & B
 \end{array}$$

$f^* = \text{Hom}_{\mathcal{C}}(f, B) :$

Exercise 10. Check that $\text{Hom}(A, -)$ and $\text{Hom}(-, B)$ are indeed functors.

We will be particularly interested in the Hom-functors in the category $R\text{-mod}$, which we will study in detail in a later chapter.

1.3 Natural transformations

Definition 1.35. Let F and G be covariant functors $\mathcal{C} \rightarrow \mathcal{D}$. A **natural transformation** between F and G is a mapping that to each object A in \mathcal{C} assigns an arrow $\eta_A \in \text{Hom}_{\mathcal{D}}(F(A), G(A))$ such that for all $f \in \text{Hom}_{\mathcal{C}}(A, B)$, the diagram

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

commutes. We sometimes write

$$\mathcal{C} \begin{array}{c} \xrightarrow{F} \\ \Downarrow \eta \\ \xrightarrow{G} \end{array} \mathcal{D}$$

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

Definition 1.36. Let F and G be contravariant functors $\mathcal{C} \rightarrow \mathcal{D}$. A **natural transformation** between F and G is a mapping that to each object A in \mathcal{C} assigns an arrow $\eta_A \in \text{Hom}_{\mathcal{D}}(F(A), G(A))$ such that for all $f \in \text{Hom}_{\mathcal{C}}(A, B)$, the diagram

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

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 1.37. Recall the abelianization functor we discussed in Example 1.27. The abelianization comes equipped with a natural projection map $\pi_G: G \rightarrow G^{\text{ab}}$, 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 $\text{ab}: \mathbf{Grp} \rightarrow \mathbf{Grp}$. On objects, the abelianizations functor is defined as $G \mapsto G^{\text{ab}}$. 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:

$$\begin{array}{ccc} G & \xrightarrow{\pi_G} & G^{\text{ab}} \\ f \downarrow & & \downarrow f^{\text{ab}} \\ H & \xrightarrow{\pi_H} & H^{\text{ab}} \end{array} .$$

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

$$\text{Grp} \begin{array}{c} \xrightarrow{\text{id}} \\ \Downarrow \pi \\ \xrightarrow{\text{ab}} \end{array} \text{Grp}.$$

Example 1.38. The determinant gives rise to a natural transformation. Fix an integer $n \geq 1$, and consider the GL_n functor

$$\text{GL}_n \text{ Ring} \rightarrow \text{Grp}$$

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 $f: R \rightarrow S$ to the map

$$\text{GL}_n(f): \text{GL}_n(R) \rightarrow \text{GL}_n(S)$$

that applies f to all the entries of each matrix $A \in \text{GL}_n(R)$, 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 Example 1.27. First, note that the determinant of an invertible matrix is a unit, so the determinant gives a map $\text{GL}_n(R) \rightarrow R^*$. Moreover, given any ring homomorphism $f: R \rightarrow S$, we have a commutative diagram

$$\begin{array}{ccc} \text{GL}_n(R) & \xrightarrow{\det} & R^* \\ f \downarrow & & \downarrow f \\ \text{GL}_n(S) & \xrightarrow{\det} & S^* \end{array}$$

Above we identified f with both the map $\text{GL}_n(f)$ 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 1.39. A **natural isomorphism** is a natural transformation η where each η_A is an isomorphism.

Exercise 11. Show that a natural transformation $\eta: F \Rightarrow G$ is a natural isomorphism if and only if there exists a natural transformation $\mu: G \Rightarrow F$ 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 with S , but no such bijection can be natural. Details below.

Exercise 12. Let \mathbf{Set}^∞ be the full subcategory of \mathbf{Set} consisting of all infinite sets. Let

$$F: \mathbf{Set}^\infty \rightarrow \mathbf{Set}^\infty$$

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 1.40. Let $F, G: \mathcal{C} \rightarrow \mathcal{D}$ be two functors between the categories \mathcal{C} and \mathcal{D} . We write

$$\mathrm{Nat}(F, G) = \{\text{natural transformations } F \Rightarrow G\}.$$

Given two categories \mathcal{C} and \mathcal{D} , one can build a **functor category**⁴ with objects all covariant functors $\mathcal{C} \rightarrow \mathcal{D}$, and arrows the corresponding natural transformations. This category is denoted $\mathcal{D}^{\mathcal{C}}$. 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 1.41. Consider natural transformations

$$\begin{array}{ccc} \mathcal{C} & \begin{array}{c} \xrightarrow{F} \\ \Downarrow \varphi \\ \xrightarrow{G} \end{array} & \mathcal{D} \end{array} \quad \text{and} \quad \begin{array}{ccc} \mathcal{C} & \begin{array}{c} \xrightarrow{G} \\ \Downarrow \eta \\ \xrightarrow{H} \end{array} & \mathcal{D}. \end{array}$$

We can compose them for form a new natural transformation

$$\begin{array}{ccc} \mathcal{C} & \begin{array}{c} \xrightarrow{F} \\ \Downarrow \eta\varphi \\ \xrightarrow{H} \end{array} & \mathcal{D}. \end{array}$$

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

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

This makes the diagram

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

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

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

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

⁴Yes, the madness is neverending.

Definition 1.42. Two categories \mathcal{C} and \mathcal{D} are **equivalent** if there exist functors $F: \mathcal{C} \rightarrow \mathcal{D}$ and $G: \mathcal{D} \rightarrow \mathcal{C}$ and two natural isomorphisms $\alpha: GF \Rightarrow 1_{\mathcal{C}}$ and $\beta: FG \Rightarrow 1_{\mathcal{D}}$. We say that a functor $F: \mathcal{C} \rightarrow \mathcal{D}$ 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 \mathcal{C} and the objects of \mathcal{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 13. Let \mathcal{C} be the category with one object C and a unique arrow 1_C . Let \mathcal{D} be the category with two objects D_1 and D_2 and four arrows: the identities 1_{D_i} and two isomorphisms $\alpha: D_1 \rightarrow D_2$ and $\beta: D_2 \rightarrow D_1$. Let \mathcal{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 \mathcal{C} and \mathcal{D} are equivalent categories.
- b) Show that \mathcal{C} and \mathcal{E} are not equivalent categories.

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

Definition 1.43. A covariant functor $F: \mathcal{C} \rightarrow \mathbf{Set}$ is **representable** if there exists an object A in \mathcal{C} such that F is naturally isomorphic to $\text{Hom}_{\mathcal{C}}(A, -)$. A contravariant functor $F: \mathcal{C} \rightarrow \mathbf{Set}$ is **representable** if there exists an object B in \mathcal{C} such that F is naturally isomorphic to $\text{Hom}_{\mathcal{C}}(-, B)$.

Example 1.44. We claim that the identity functor $\mathbf{Set} \rightarrow \mathbf{Set}$ is representable. Let $\mathbf{1}$ be a singleton set. Given any set X , there is a bijection between elements $x \in X$ and functions $\mathbf{1} \rightarrow X$ sending the one element in $\mathbf{1}$ to each x . Moreover, given any other set Y , and a function $f: X \rightarrow Y$, our bijections make the following diagram commute:

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

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

Exercise 14. Show that the forgetful functor $\mathbf{Grp} \rightarrow \mathbf{Set}$ is representable.

Exercise 15. Given a ring R , show that the forgetful functor $R\text{-mod} \rightarrow \mathbf{Set}$ is representable.

The Yoneda Lemma tells us that in order to study a locally small category \mathcal{C} , it is in many ways sufficient to study the category of functors from \mathcal{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 1.45 (Yoneda Lemma). *Let \mathcal{C} be a locally small category, and fix an object A in \mathcal{C} . Let $F: \mathcal{C} \rightarrow \mathbf{Set}$ be a covariant functor. Then there is a bijection*

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

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

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

$$\begin{array}{ccc}
 \mathrm{Hom}_{\mathcal{C}}(A, A) & \xrightarrow{\mathrm{Hom}_{\mathcal{C}}(A, f)} & \mathrm{Hom}_{\mathcal{C}}(A, X) \\
 \downarrow \varphi_A & & \downarrow \varphi_X \\
 & \begin{array}{ccc} 1_A & \xrightarrow{\quad} & f \\ \downarrow & & \downarrow \\ u & \xrightarrow{\quad} & (F(f))u = \varphi_X(f) \end{array} & \\
 F(A) & \xrightarrow{F(f)} & F(X)
 \end{array}$$

Our bijection will be defined by

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

We should first check that this makes sense: arrows in \mathbf{Set} are just functions between sets, and so φ_A is a function of sets $\mathrm{Hom}_{\mathcal{C}}(A, A) \rightarrow F(A)$. Also, $\mathrm{Hom}_{\mathcal{C}}(A, A)$ 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 $f \in \mathrm{Hom}_{\mathcal{C}}(A, X)$, 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 $\varphi_X \mathrm{Hom}_{\mathcal{C}}(A, f)$ 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 \mathcal{C} and any arrow $f \in \mathrm{Hom}_{\mathcal{C}}(A, X)$, 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 $f \in \mathrm{Hom}_{\mathcal{C}}(A, X)$ 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 $g: X \rightarrow Y$, the diagram

$$\begin{array}{ccc} \text{Hom}_{\mathcal{C}}(A, X) & \xrightarrow{\varphi_X} & F(X) \\ g_* \downarrow & & \downarrow F(g) \\ \text{Hom}_{\mathcal{C}}(A, Y) & \xrightarrow{\varphi_Y} & F(Y) \end{array}$$

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

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

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 $\eta: F \rightarrow G$, the following diagram must commute:

$$\begin{array}{ccc} \text{Nat}(\text{Hom}_{\mathcal{C}}(A, -), F) & \xrightarrow{\gamma_F} & F(A) \\ \eta_* \downarrow & & \downarrow \eta_A \\ \text{Nat}(\text{Hom}_{\mathcal{C}}(A, -), G) & \xrightarrow{\gamma_G} & G(A) \end{array}$$

Given a natural transformation φ between $\text{Hom}_{\mathcal{C}}(A, -)$ 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 $f: A \rightarrow B$, the diagram

$$\begin{array}{ccc} \text{Nat}(\text{Hom}_{\mathcal{C}}(A, -), F) & \xrightarrow{\gamma_A} & F(A) \\ (f^*)^* \downarrow & & \downarrow F(f) \\ \text{Nat}(\text{Hom}_{\mathcal{C}}(B, -), F) & \xrightarrow{\gamma_B} & F(B) \end{array}$$

commutes. Given a natural transformation φ between $\text{Hom}_{\mathcal{C}}(A, -)$ 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

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

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. \square

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

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

which is natural on both B and F .

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

Remark 1.46. Notice that the Yoneda Lemma says in particular that the collection of all natural transformations from $\mathrm{Hom}_{\mathcal{C}}(A, -)$ to F is a set. This wasn't clear a priori, since the collection of objects in \mathcal{C} is not necessarily a set.

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

Remark 1.47. If we apply the [Yoneda Lemma](#) to the case when F itself is also a Hom functor, say $F = \mathrm{Hom}_{\mathcal{C}}(B, -)$, the Yoneda Lemma says that there is a bijection between $\mathrm{Nat}(\mathrm{Hom}_{\mathcal{C}}(A, -), \mathrm{Hom}_{\mathcal{C}}(B, -))$ and $\mathrm{Hom}_{\mathcal{C}}(B, A)$. In particular, each arrow in \mathcal{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 \mathcal{C} to \mathbf{Set} . It is common to refer to both [Theorem 1.45](#) and [Theorem 1.49](#) as the Yoneda Lemma.

Remark 1.48. Let \mathcal{C} be a locally small category. Each arrow $f: A \rightarrow B$ in \mathcal{C} gives rise to a natural transformation $\mathrm{Hom}_{\mathcal{C}}(-, A) \Rightarrow \mathrm{Hom}_{\mathcal{C}}(-, B)$ that sends each object X to the arrow (function)

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

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

$$\begin{array}{ccccc}
 X & & \text{Hom}_{\mathcal{C}}(X, A) & \xrightarrow{f^*} & \text{Hom}_{\mathcal{C}}(X, B) \\
 \downarrow g & & \uparrow \text{Hom}_{\mathcal{C}}(g, A)=g^* & & \uparrow \text{Hom}_{\mathcal{C}}(g, B)=g^* \\
 Y & & \text{Hom}_{\mathcal{C}}(Y, A) & \xrightarrow{f_*} & \text{Hom}_{\mathcal{C}}(Y, B)
 \end{array}$$

This diagram commutes since

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

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

$$\begin{array}{ccc}
 \text{Hom}_{\mathcal{C}}(B, X) & \xrightarrow{f^*} & \text{Hom}_{\mathcal{C}}(A, X) \\
 g \vdash & & gf.
 \end{array}$$

Theorem 1.49 (Yoneda Embedding). *Let \mathcal{C} be a locally small category. The covariant functor*

$$\begin{array}{ccc}
 \mathcal{C} & \longrightarrow & \mathbf{Set}^{\mathcal{C}^{op}} \\
 A & & \text{Hom}_{\mathcal{C}}(-, A) \\
 f \downarrow & \longmapsto & \downarrow f_* \\
 B & & \text{Hom}_{\mathcal{C}}(-, B)
 \end{array}$$

from \mathcal{C} to the category of contravariant functors $\mathcal{C} \rightarrow \mathbf{Set}$ is an embedding. Moreover, the contravariant functor

$$\begin{array}{ccc}
 \mathcal{C} & \longrightarrow & \mathbf{Set}^{\mathcal{C}} \\
 A & & \text{Hom}_{\mathcal{C}}(A, -) \\
 f \downarrow & \longmapsto & \uparrow f^* \\
 B & & \text{Hom}_{\mathcal{C}}(B, -)
 \end{array}$$

from the category \mathcal{C} to the category of covariant functors $\mathcal{C} \rightarrow \mathbf{Set}$ 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 \mathcal{C} , we have bijections

$$\text{Hom}_{\mathcal{C}}(A, B) \cong \text{Nat}(\text{Hom}_{\mathcal{C}}(-, A), \text{Hom}_{\mathcal{C}}(-, B))$$

and

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

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

This follows from Remark 1.47, but let's carefully check the details. First, in Remark 1.48 we have already checked that each arrow is indeed taken to a natural transformation, so we just need to check injectivity and surjectivity at the level of arrows.

The Yoneda Lemma applied here tells us that each natural transformation φ between $\text{Hom}_{\mathcal{C}}(B, -)$ and $F = \text{Hom}_{\mathcal{C}}(A, -)$ corresponds to an element $u \in F(B) = \text{Hom}_{\mathcal{C}}(A, B)$, 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 \mathcal{C} has $\varphi_X: \text{Hom}_{\mathcal{C}}(B, X) \rightarrow \text{Hom}_{\mathcal{C}}(A, X)$ given by

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

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. \square

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

Proof. The Yoneda Embeddings from Theorem 1.49 are fully faithful, and thus by Exercise 9 they must reflect isomorphisms. A natural isomorphism between the functors $\text{Hom}_{\mathcal{C}}(X, -)$ and $\text{Hom}_{\mathcal{C}}(Y, -)$ (or the functors $\text{Hom}_{\mathcal{C}}(-, X)$ and $\text{Hom}_{\mathcal{C}}(-, Y)$) is an isomorphism in the target functor category, and it corresponds to f_* (respectively, f^*) for some arrow f from Y to X . By Exercise 9, f must be an isomorphism. In particular, X and Y are isomorphic. \square

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

- To give a natural transformation from $\text{Hom}(A, -)$ to F is the same as giving an element in the set $F(A)$.
- The collection of all natural transformations from $\text{Hom}(A, -)$ 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 \mathcal{C} can be embedded into the functor category of (covariant or contravariant) functors from \mathcal{C} to **Set**. So rather than studying the category \mathcal{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.

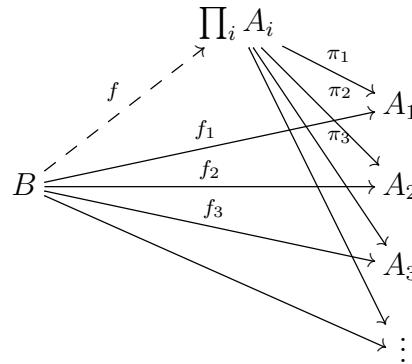
1.5 Products and coproducts

Definition 1.51. Let \mathcal{C} be a locally small category, and consider a family of objects $\{A_i\}_{i \in I}$ in \mathcal{C} . The **product** of the A_i is an object in \mathcal{C} , denoted by $\prod_i A_i$, together with arrows $\pi_j \in \text{Hom}_{\mathcal{C}}(\prod_i A_i, A_j)$ for each j , called **projections**, satisfying the following universal property: given any object B in \mathcal{C} and arrows $f_i: B \rightarrow A_i$ for each i , there exists a unique arrow f such that

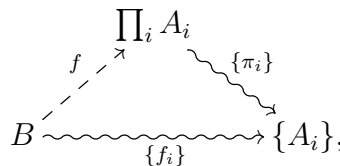
$$\begin{array}{ccc} B & & \\ \downarrow \exists! f & \searrow f_j & \\ \prod_i A_i & \xrightarrow{\pi_j} & A_j \end{array}$$

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

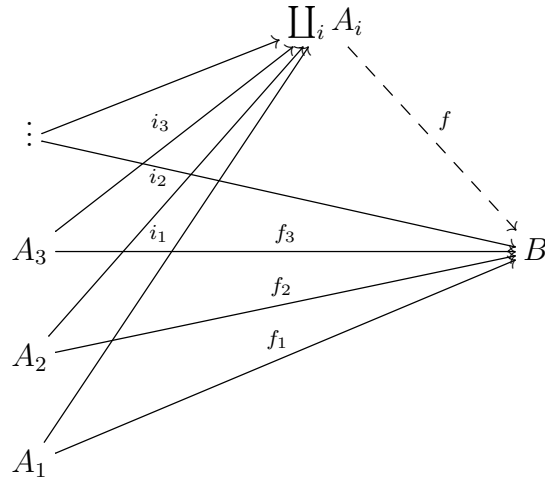
The dual notion is the coproduct.

Definition 1.52. Let \mathcal{C} be a locally small category, and consider a family of objects $\{A_i\}_{i \in I}$ in \mathcal{C} . The **coproduct** of the A_i is an object in \mathcal{C} , denoted by $\coprod_i A_i$, together with arrows $\iota_j \in \text{Hom}_{\mathcal{C}}(A_j, \coprod_i A_i)$ for each j , satisfying the following universal property: given any object B in \mathcal{C} and arrows $f_i: A_i \rightarrow B$ for each i , the following diagram commutes:

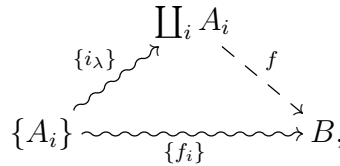
$$\begin{array}{ccc} B & & \\ \uparrow \exists! f & \swarrow f_j & \\ \coprod_i A_i & \xleftarrow{\iota_j} & A_j \end{array}$$

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

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



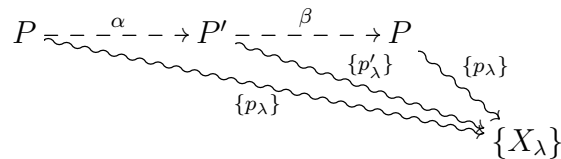
We can also take a “big picture” view of the universal property of the coproduct:



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

Theorem 1.53. *If $(P, \{p_\lambda : P \rightarrow X_\lambda\}_{\lambda \in \Lambda})$ and $(P', \{p'_\lambda : P' \rightarrow X_\lambda\}_{\lambda \in \Lambda})$ are both products for the same family of objects $\{X_\lambda\}_{\lambda \in \Lambda}$ in a category \mathcal{C} , then there is a unique isomorphism $\alpha : P \xrightarrow{\sim} P'$ 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 $\beta : P' \rightarrow P$ such that $p_\lambda \circ \beta = p'_\lambda$. Switching roles, we obtain a unique map $\alpha : P \rightarrow P'$ such that $p'_\lambda \circ \alpha = p_\lambda$.

Consider the composition $\beta \circ \alpha : P \rightarrow P$. We have $p_\lambda \circ \beta \circ \alpha = p'_\lambda \circ \alpha = p_\lambda$ for all λ . The identity map $1_P : P \rightarrow P$ 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. \square

Exercise 17. Prove the analogous statement to Theorem 1.53 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 1.54. 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 **Sgrp**, **Grp**, **Ring**, **R-Mod**, **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 1.55. *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 $\pi_\lambda: \prod_{\gamma \in \Lambda} X_\gamma \rightarrow X_\lambda$ 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 \mathcal{C} be one of these categories, and suppose that we have morphisms $g_\lambda: Y \rightarrow X_\lambda$ for all λ in \mathcal{C} . We need to show there is a unique morphism $\phi: Y \rightarrow \prod_{\lambda \in \Lambda} X_\lambda$ 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 \mathcal{C} ; we leave the details as an exercise. \square

Example 1.56. 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 1.57. 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 $g_\lambda : M_\lambda \rightarrow N$ for each λ , we need to show that there is a unique R -module homomorphism $\alpha : \bigoplus_{\lambda \in \Lambda} M_\lambda \rightarrow N$ 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. \square

Remark 1.58. 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 1.59. For any indexing set Λ , $\bigoplus_{\lambda \in \Lambda} R$ 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 1.60.

- 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 1.61. Let (I, \geq) be a partially ordered set and let \mathcal{C} be a category. An **inverse system** in \mathcal{C} indexed by I is a contravariant functor $\mathbf{PO}(I) \rightarrow \mathcal{C}$.

Remark 1.62. Let's unwrap the definition of inverse system a bit. For each $i \in I$, we get an object M_i in \mathcal{C} . Moreover, in the category $\mathbf{PO}(I)$, there is exactly one arrow $i \rightarrow j$ for each $i \leq j$, and the image of this arrow under any contravariant functor $\mathbf{PO}(I) \rightarrow \mathcal{C}$ is an arrow $M_j \rightarrow M_i$. Finally, our functor must preserve compositions of arrows, so whenever $k \geq j \geq i$, the arrow $M_k \rightarrow M_i$ should match the composition of arrows through j . Thus an inverse system in \mathcal{C} indexed by I consists of the following data:

- for each $i \in I$, an object M_i in \mathcal{C} , and
- for each $i \leq j$, an arrow $\varphi_i^j: M_j \rightarrow M_i$ in \mathcal{C}

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

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

Note moreover that $\varphi_i^i = \text{id}_{M_i}$, 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 1.63.

a) An inverse system in a category \mathcal{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 $X_j \rightarrow X_i$ 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 $L \leq N$ if and only if $L \supseteq N$. Whenever $N \subseteq L$, we have an inclusion map $N \rightarrow L$, 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 $i \leq j$ 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 $1 \leq 2$ and $1 \leq 3$, then an inverse system indexed by I is just a diagram of the form

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

Exercise 18. Let J be an ideal in a commutative ring R , and consider its n th 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 $m \geq n$, 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 $R\text{-}\mathbf{Mod}$ 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 1.64. Let \mathcal{C} be a category and let $\{M_i, \varphi_i^j\}_i$ be an inverse system on \mathcal{C} indexed by I . The **limit** or **inverse limit** of $\{M_i, \varphi_i^j\}$ consists of an object

$$\varprojlim M_i$$

and arrows

$$\pi_i: \varprojlim M_i \rightarrow M_i$$

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

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

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

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

commutes, there exists a unique arrow $f: X \rightarrow \varprojlim M_i$ such that

$$\begin{array}{ccc} \varprojlim M_i & \xleftarrow{\exists! f} & X \\ \pi_j \searrow & & \swarrow f_j \\ & M_j & \end{array}$$

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.

Remark 1.65. Given an inverse system $\{M_i, \varphi_i^j\}$ indexed by I in a category \mathcal{C} , say corresponding to the contravariant functor $\varphi : I \rightarrow \mathcal{C}$, suppose that its limit exists, and let $L = \varprojlim M_i$. The projections π_i give us commutative diagrams

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

This is the same data as a natural transformation

$$\mathbf{PO}(I) \begin{array}{c} \xrightarrow{\Delta L} \\ \Downarrow \\ \xrightarrow{\varphi} \end{array} \mathcal{C}.$$

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 1.66. 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 $X \rightarrow L$.

Exercise 19. 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 1.67. *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 $\varphi_i^j : M_j \rightarrow M_i$. Let

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

One can show (exercise!) that this is a submodule of the product of the M_i . For each i , let $\pi_i : L \rightarrow M_i$ be the restriction of the projection maps $\prod M_i \rightarrow M_i$ 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 $f_i : X \rightarrow M_i$ such that $\varphi_i^j f_j = f_i$ for all $i \leq j$. Define

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

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 $f: X \rightarrow L$. 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 $f: X \rightarrow L$ 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

$$\begin{array}{ccc} \varprojlim M_i & \xleftarrow{\quad f \quad} & X \\ & \searrow \pi_i \quad \swarrow f_i & \\ & M_i & \end{array}$$

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 $X \rightarrow L$ such that

$$\begin{array}{ccc} \varprojlim M_i & \xleftarrow{\quad h \quad} & X \\ & \searrow \pi_i \quad \swarrow f_i & \\ & M_i & \end{array}$$

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. \square

Remark 1.68. One can adapt the proof of Theorem 1.67 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

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

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

Example 1.69.

- 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

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

is the J -adic completion of R .

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

Definition 1.70. Let (I, \geq) be a partially ordered set and let \mathcal{C} be a category. A **direct system** in \mathcal{C} indexed by I is a covariant functor $\mathbf{PO}(I) \rightarrow \mathcal{C}$.

Remark 1.71. An inverse system in \mathcal{C} indexed by I consists of the following data:

- for each $i \in I$, an object M_i in \mathcal{C} , and
- for each $i \leq j$, an arrow $\varphi_j^i: M_i \rightarrow M_j$ in \mathcal{C}

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

$$\begin{array}{ccc} M_i & \xrightarrow{\varphi_i^k} & M_k \\ & \searrow \varphi_k^i & \nearrow \varphi_k^j \\ & M_j & \end{array}$$

Note moreover that $\varphi_i^i = \text{id}_{M_i}$, 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 1.72.

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

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

All the other arrows $X_i \rightarrow X_j$ 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 $N \rightarrow L$, 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 $1 \leq 2$ and $1 \leq 3$, then a direct system indexed by I is just a diagram of the form

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

Definition 1.73. Let \mathcal{C} be a category and let $\{M_i, \varphi_j^i\}_i$ be a direct system on \mathcal{C} indexed by I . The **colimit** or **direct limit** of $\{M_i, \varphi_j^i\}$ consists of an object

$$\varinjlim M_i$$

and arrows

$$\alpha_i: M_i \rightarrow \varinjlim M_i$$

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 $f_i: M_i \rightarrow X$ such that $f_j \varphi_j^i = f_i$ for all i, j , meaning that the diagram

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

commutes, there exists a unique arrow $f: \varinjlim M_i \rightarrow X$ such that

$$\begin{array}{ccc} \varinjlim M_i & \xrightarrow{\exists! f} & X \\ \alpha_j \swarrow & & \nearrow f_j \\ & M_j & \end{array}$$

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.

Remark 1.74. Given a direct system $\{M_i, \varphi_j^i\}$ indexed by I in a category \mathcal{C} , say corresponding to the covariant functor $\varphi: I \rightarrow \mathcal{C}$, suppose that its colimit exists, and let $L = \varinjlim M_i$. The α_i give us commutative diagrams

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

This is the same data as a natural transformation

$$\begin{array}{ccc} \mathcal{C} & \xrightarrow{\varphi} & \mathcal{C} \\ & \Downarrow & \\ & \Delta L & \end{array}.$$

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 1.75. 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 $C \rightarrow X$.

Exercise 20. 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 1.76. *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 $\varphi_j^i: M_j \rightarrow M_i$. Let $\iota_i: M_i \rightarrow \bigoplus_j M_j$ be the inclusions into the direct sum, let S be the submodule of $\bigoplus M_i$ generated by all elements of the form

$$\iota_i(\varphi_j^i(m_i)) - \iota_i(m_i),$$

and define

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

For each i , let

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

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

Remark 1.77. One can adapt the proof of Theorem 1.67 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:

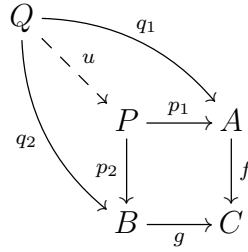
Definition 1.78. Let \mathcal{C} be a category. A **pullback** of the arrows f and g consists of an object P and arrows p_1 and p_2 such that

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

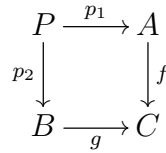
commutes, and satisfying the following universal property: for all objects Q and arrows q_1 and q_2 such that

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

commutes, there exists a unique u such that

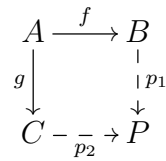


commutes. One sometimes refers to the following diagram as a **pullback diagram**:

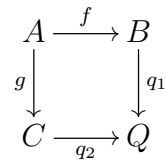


The dual construction is the pushout.

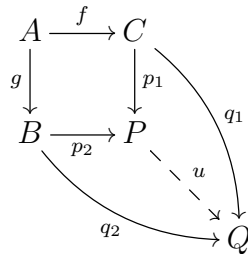
Definition 1.79. Let \mathcal{C} be a category. A **pushout** of the arrows f and g consists of an object P and arrows p_1 and p_2 such that



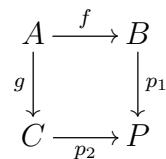
commutes, and satisfying the following universal property: for all objects Q and arrows q_1 and q_2 such that



commutes, there exists a unique u such that



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



Exercise 21. 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 1.67 and Theorem 1.76 that $R\text{-Mod}$ has all limits and colimits. In the case of pullbacks and pushouts, one can describe the corresponding module in a more manageable way.

Exercise 22. Explicitly describe pullbacks and pushouts in $R\text{-Mod}$.

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 1.80. Let \mathcal{C} be a category and let J be a small category. A **diagram** in \mathcal{C} of **shape** J is a functor $J \rightarrow \mathcal{C}$. We may call J the **index category**.

Remark 1.81. Let \mathcal{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 $J \rightarrow \mathcal{C}$ is to give

- a set $\{X_i\}_{i \in I}$ in \mathcal{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 $\text{Hom}_{\mathcal{C}}(X_i, X_j)$ indexed by the set $\text{Hom}_I(i, j)$

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 $F: J \rightarrow \mathcal{C}$, 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 1.82. Consider a diagram in \mathcal{C} with objects $\{X_i\}_{i \in I}$ in \mathcal{C} and arrows $A_{i,j} = \{f_\alpha\}$. A **cone** over this diagram consists of

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

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

$$\begin{array}{ccc} & C & \\ p_i \swarrow & & \searrow p_j \\ X_i & \xrightarrow{f} & X_j \end{array}$$

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

- an object C in \mathcal{C} , and

- for each $i \in I$, an arrow $p_i: X_i \rightarrow C$

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

$$\begin{array}{ccc} X_i & \xrightarrow{f} & X_j \\ & \searrow p_i & \swarrow p_j \\ & C & \end{array}$$

Definition 1.83. Consider a diagram in \mathcal{C} with objects $\{X_i\}_{i \in I}$ in \mathcal{C} and arrows $A_{i,j} = \{f_\alpha\}$. The **limit** of this diagram is, if it exists, a cone

$$\begin{array}{ccc} & \lim X_i & \\ p_i \swarrow & & \searrow p_j \\ X_i & \xrightarrow{\quad} & X_j \end{array}$$

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

$$\begin{array}{ccc} & C & \\ q_i \swarrow & & \searrow q_j \\ X_i & \xrightarrow{f} & X_j \end{array}$$

there exists a unique arrow $u: C \rightarrow \lim X_i$ such that

$$\begin{array}{ccc} & \lim X_i & \\ u \swarrow & & \searrow p_j \\ C & \xrightarrow{q_j} & X_j \end{array}$$

commutes.

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

$$\begin{array}{ccc} X_i & \xrightarrow{\quad} & X_j \\ & \searrow p_i & \swarrow p_j \\ & \text{colim } X_i & \end{array}$$

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

$$\begin{array}{ccc} X_i & \xrightarrow{\quad} & X_j \\ & \searrow q_i & \swarrow q_j \\ & C & \end{array}$$

there exists a unique arrow $u: C \rightarrow \text{colim } X_i$ such that

$$\begin{array}{ccc} X_j & \xrightarrow{q_j} & C \\ & \searrow p_j & \swarrow f \\ & \text{colim } X_i & \end{array}$$

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 1.84. A covariant functor $F: \mathcal{C} \rightarrow \mathcal{D}$

- **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 $\alpha_i: M_i \rightarrow \operatorname{colim} M_i$ form the colimit of diagram D , then $F(\operatorname{colim} M_i)$ is the colimit of the diagram $F \circ D$ with insertion arrows $F(\alpha_j): F(M_j) \rightarrow F(\operatorname{colim} M_i)$.

- **preserves limits** if

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

More precisely, if $\operatorname{lim} M_i$ is the limit of a diagram D with projections $\pi_j: \operatorname{lim} M_i \rightarrow M_j$, then the object $F(\operatorname{lim} M_i)$ and the projection arrows $F(\pi_j): F(\operatorname{lim} M_i) \rightarrow F(M_j)$ form a limit of the diagram $F \circ D$.

Definition 1.85. A contravariant functor $F: \mathcal{C} \rightarrow \mathcal{D}$ **converts limits to colimits** or **sends limits to colimits** if

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

Similarly, a contravariant functor $F: \mathcal{C} \rightarrow \mathcal{D}$ **converts colimits to limits** or **sends colimits to limits** if

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

The Hom functors preserve limits and colimits.

Theorem 1.86. *Let \mathcal{C} be any category and let A be an object in \mathcal{C} .*

- a) *If the limit $\operatorname{lim}_i M_i$ exists, then there is a natural isomorphism*

$$\operatorname{Hom}_{\mathcal{C}}(A, \operatorname{lim}_i M_i) \cong \operatorname{lim}_i \operatorname{Hom}_{\mathcal{C}}(A, M_i).$$

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

- b) *If the limit $\operatorname{lim}_i M_i$ exists, then there is a natural isomorphism*

$$\operatorname{Hom}_{\mathcal{C}}(\operatorname{colim}_i M_i, A) \cong \operatorname{lim}_i \operatorname{Hom}_{\mathcal{C}}(M_i, A).$$

In particular, the limit of $\operatorname{Hom}_{\mathcal{C}}(M_i, A)$ 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 Example 1.44, 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. 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 1.87. Let \mathcal{C} be a locally small category. A **universal property** of an object C in \mathcal{C} consists of a representable functor $F: \mathcal{C} \rightarrow \mathbf{Set}$ together with a **universal element** $X \in F(C)$ such that F is naturally isomorphic to either $\mathrm{Hom}_{\mathcal{C}}(C, -)$ (if F is covariant) or $\mathrm{Hom}_{\mathcal{C}}(-, C)$ (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 1.88. Let $F: \mathcal{C} \rightarrow \mathcal{D}$ be covariant functor. Given an object $D \in \mathcal{D}$, a **universal arrow from D to F** is a unique pair (U, u) where U is an object in \mathcal{C} and a unique arrow $u \in \mathrm{Hom}_{\mathcal{D}}(D, F(U))$ with the following **universal property**: for any arrow $f \in \mathrm{Hom}_{\mathcal{D}}(D, F(Y))$, there exists a unique arrow $h \in \mathrm{Hom}_{\mathcal{C}}(U, Y)$ such that the following diagram commutes:

$$\begin{array}{ccc} U & & D \xrightarrow{u} F(U) \\ \downarrow h & & \searrow f \quad \downarrow F(h) \\ Y & & F(Y). \end{array}$$

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

$$\begin{array}{ccc} U & & D \xleftarrow{u} F(U) \\ \uparrow h & & \swarrow f \quad \uparrow F(h) \\ Y & & F(Y). \end{array}$$

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

Remark 1.89. Let $F: \mathcal{C} \rightarrow \mathcal{D}$ be a covariant functor, and fix an object U in \mathcal{C} , an object D in \mathcal{D} , and an arrow $u \in \text{Hom}_{\mathcal{D}}(D, F(U))$. Notice that $\text{Hom}_{\mathcal{D}}(D, F(-))$ determines a covariant functor $\mathcal{C} \rightarrow \mathbf{Set}$. By the [Yoneda Lemma](#), the following is a recipe for a natural transformation between $\text{Hom}_{\mathcal{C}}(U, -)$ and $\text{Hom}_{\mathcal{D}}(D, F(-))$: for each object Y in \mathcal{C} and each arrow $h \in \text{Hom}_{\mathcal{C}}(U, Y)$, set

$$\varphi_X(h) := \text{Hom}_{\mathcal{D}}(D, F(h))(u).$$

Notice that

$$\begin{array}{ccc} \text{Hom}_{\mathcal{D}}(D, F(U)) & \xrightarrow{\text{Hom}_{\mathcal{D}}(D, F(h))} & \text{Hom}_{\mathcal{D}}(D, F(Y)) , \\ f \mapsto & & F(h) \circ u \end{array}$$

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

We get the following commutative diagram:

$$\begin{array}{ccc} \text{Hom}_{\mathcal{C}}(U, U) & \xrightarrow{\text{Hom}_{\mathcal{C}}(U, h)} & \text{Hom}_{\mathcal{C}}(U, Y) \\ \downarrow \varphi_U & & \downarrow \varphi_Y \\ \begin{array}{ccc} 1_U & \xrightarrow{\quad} & h \\ \downarrow & & \downarrow \\ u & \xrightarrow{\quad} & F(h) \circ u =: \varphi_Y(h) \end{array} \\ \text{Hom}_{\mathcal{D}}(D, F(U)) & \xrightarrow{\text{Hom}_{\mathcal{D}}(D, F(h))} & \text{Hom}_{\mathcal{D}}(D, F(Y)) \end{array}$$

Given an arrow $f \in \text{Hom}_{\mathcal{D}}(D, F(Y))$, $\varphi_Y(h) = f$ for some $h \in \text{Hom}_{\mathcal{C}}(U, Y)$ 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 \mathcal{C} and every $f \in \text{Hom}_{\mathcal{D}}(D, F(Y))$ there exists a unique $h \in \text{Hom}_{\mathcal{C}}(U, Y)$ 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 $\text{Hom}_{\mathcal{D}}(D, F(-)): \mathcal{C} \rightarrow \mathbf{Set}$, via $u \in \text{Hom}_{\mathcal{D}}(D, F(U))$.

Similarly, one can prove the dual equivalence:

- (U, u) is a universal arrow from F to D .
- U represents the functor $\text{Hom}_{\mathcal{D}}(F(-), D): \mathcal{C} \rightarrow \mathbf{Set}$, via $u \in \text{Hom}_{\mathcal{D}}(F(U), D)$.

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

Remark 1.90. 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 \mathcal{C} . To do that, we need to consider the **product category** $\mathcal{C} \times \mathcal{C}$ with objects given by pairs (C_1, C_2) of objects in \mathcal{C} and arrows in $(C_1, C_2) \rightarrow (C_3, C_4)$ given by pairs of arrows (f_1, f_2) with $f_1 \in \text{Hom}_{\mathcal{C}}(C_1, C_3)$ and $f_2 \in \text{Hom}_{\mathcal{C}}(C_2, C_4)$. The diagonal functor $\Delta: \mathcal{C} \rightarrow \mathcal{C} \times \mathcal{C}$ is exactly what it sounds like: $\Delta(C) = (C, C)$ for every object C in \mathcal{C} and $\Delta(f) = (f, f)$ for every arrow f in \mathcal{C} .

Given objects X and Y in \mathcal{C} , consider the projection arrows $\pi_1: X \times Y \rightarrow X$ and $\pi_2: X \times Y \rightarrow Y$. We claim that the object $X \times Y$ together with the arrow (π_1, π_2) in $\mathcal{C} \times \mathcal{C}$ form a universal arrow from Δ to (X, Y) in $\mathcal{C} \times \mathcal{C}$. Why? This means that given any arrow $(f_1, f_2) \in \text{Hom}_{\mathcal{C} \times \mathcal{C}}(\Delta(X_1 \times X_2), \Delta(Y))$, there exists a unique $h \in \mathcal{C}(X_1 \times X_2, Y)$ such that

$$\begin{array}{ccc} X_1 \times X_2 & & (X_1, X_2) \xleftarrow{(\pi_1, \pi_2)} \Delta(X_1 \times X_2) \\ \uparrow h & \swarrow (f_1, f_2) & \uparrow \Delta(h) \\ Y & & \Delta(Y) \end{array}$$

commutes. This is indeed the universal property of products we just described less formally above: given $f_1: Y \rightarrow X_1$ and $f_2: Y \rightarrow X_2$, there is a unique $h: Y \rightarrow X_1 \times X_2$ such that

$$\begin{array}{ccc} X_1 \times X_2 & & (X_1, X_2) \xleftarrow{(\pi_1, \pi_2)} (X_1 \times X_2, X_1 \times X_2) \\ \uparrow h & \swarrow (f_1, f_2) & \uparrow \Delta(h) \\ Y & & (Y, Y) \end{array}$$

Equivalently, following the recipe we described in Remark 1.89, the universal property of the product is encoded in the representable functor $\text{Hom}_{\mathcal{C} \times \mathcal{C}}(\Delta(-), (X_1, X_2))$, which is represented by $X_1 \times X_2$ via (π_1, π_2) . So more precisely, that says that there is a natural isomorphism

$$\text{Hom}_{\mathcal{C}}(-, X_1 \times X_2) \cong \text{Hom}_{\mathcal{C} \times \mathcal{C}}(\Delta(-), (X_1, X_2)),$$

and that is precisely the natural transformation that the Yoneda bijection we constructed in Theorem 1.45 takes to $(\pi_1, \pi_2) \in \text{Hom}_{\mathcal{C}}(\Delta(X_1 \times X_2), (X_1, X_2))$. If we follow that bijection, our natural isomorphism φ sends an object Y in \mathcal{C} to the arrow

$$\begin{aligned} \text{Hom}_{\mathcal{C}}(Y, X_1 \times X_2) &\xrightarrow{\varphi_Y} \text{Hom}_{\mathcal{C} \times \mathcal{C}}(\Delta(Y), (X_1, X_2)) \\ f &\longmapsto \left(\Delta(Y) \xrightarrow{(f, f)} \Delta(X_1 \times X_2) \xrightarrow{(\pi_1, \pi_2)} (X_1, X_2) \right). \end{aligned}$$

Since φ_Y is a bijection, every arrow $(f_1, f_2) \in \text{Hom}_{\mathcal{C} \times \mathcal{C}}(\Delta(Y), (X_1, X_2))$ is $\varphi_Y(f)$ for some $f \in \text{Hom}_{\mathcal{C}}(Y, X_1 \times X_2)$. Ultimately, this means that there exists f 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!

1.8 Adjoint functors

Universal properties are closely related to adjoint functors.

Definition 1.91. Let \mathcal{C} and \mathcal{D} be locally small categories. Two covariant functors

$$\mathcal{C} \begin{matrix} \xrightarrow{F} \\ \xleftarrow{G} \end{matrix} \mathcal{D}$$

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

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

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

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

commute for all $C \in \mathcal{C}$ and all $D \in \mathcal{D}$. 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 1.92. 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 **Free**: **Set** \rightarrow R -**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 -**Mod** \rightarrow **Set**. That is, there is a natural bijection

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

(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 $\bigoplus_I R \rightarrow M$ 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 1.93. We can rephrase the condition that $G: \mathcal{D} \rightarrow \mathcal{C}$ has a left adjoint functor $F: \mathcal{C} \rightarrow \mathcal{D}$ as follows: for every object C in \mathcal{C} , there is a universal arrow from C to G , and for every object D in \mathcal{D} there exists a universal arrow from F to D . To see that, let $\eta_D \in \text{Hom}_{\mathcal{D}}(F(G(D)), D)$ be the image of the identity on $\text{Hom}_{\mathcal{D}}(G(D), G(D))$ via the bijection

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

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

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

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

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

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

$$\begin{array}{ccc} 1_{F(C)} & \xrightarrow{\quad} & \varepsilon_C \\ \downarrow & \text{Hom}_{\mathcal{D}}(F(C), F(C)) \xrightarrow{\cong} \text{Hom}_{\mathcal{C}}(C, GF(C)) & \downarrow \\ & \downarrow h_* & \downarrow G(h)_* \\ & \text{Hom}_{\mathcal{D}}(F(C), Y) \xrightarrow{\cong} \text{Hom}_{\mathcal{C}}(C, G(Y)) & \\ \downarrow h & & \downarrow f \end{array}$$

On the one hand, such an h does exist: just take $h \in \text{Hom}_{\mathcal{C}}(F(C), Y)$ that is sent to f via the bijection between $\text{Hom}_{\mathcal{D}}(F(C), Y)$ and $\text{Hom}_{\mathcal{C}}(C, G(Y))$. 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 $f \in \text{Hom}_{\mathcal{D}}(F(Y), D)$, there exists a unique $h \in \text{Hom}_{\mathcal{C}}(Y, G(D))$ such that the following diagram commutes:

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

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

$$\begin{array}{ccc}
 1_{G(D)} & \xrightarrow{\quad} & \eta_D \\
 \downarrow h & \begin{array}{ccc} \text{Hom}_{\mathcal{C}}(G(D), G(D)) & \xrightarrow{\cong} & \text{Hom}_{\mathcal{D}}(D, FG(D)) \\ \downarrow h^* & & \downarrow F(h)^* \\ \text{Hom}_{\mathcal{C}}(G(D), Y) & \xrightarrow{\cong} & \text{Hom}_{\mathcal{D}}(D, F(Y)) \end{array} & \downarrow f
 \end{array}$$

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

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

Exercise 23. 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 1.94. First, we view \mathbb{Z} as a partially ordered set under \geq . As in Example 1.8, \mathbb{Z} now gives us a category whose objects are the integers, and where we have an arrow in $\text{Hom}_{\mathbb{Z}}(n, m)$ if $n \geq m$. If we ignore the identity maps $\text{Hom}_{\mathbb{Z}}(n, n)$ 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 $F: \mathbb{Z} \rightarrow \mathbf{Ab}$: 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 $(F_\bullet, \partial_\bullet^F)$ and $(G_\bullet, \partial_\bullet^G)$ be complexes. A **map of complexes** or a **chain map**, which we write as $h: (F_\bullet, \partial_\bullet^F) \rightarrow (G_\bullet, \partial_\bullet^G)$ or simply $h: F \rightarrow G$, is a sequence of homomorphisms of R -modules $h_n: F_n \rightarrow G_n$ such that the following diagram commutes:

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

This means that $h_n \partial_{n+1}^F = \partial_{n+1}^G h_{n+1}$ for all n .

Note that throughout, whenever we call a function $f: M \rightarrow N$ 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 $(F_\bullet, \partial_\bullet) \rightarrow (F_\bullet, \partial_\bullet)$ 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 $\text{Ch}(R\text{-mod})$ or simply $\text{Ch}(R)$, 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 $\text{Ch}(\mathbf{Ab})$ for $\text{Ch}(\mathbb{Z})$, 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 24. Show that the isomorphisms in the category $\text{Ch}(R)$ are precisely the maps of complexes

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

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 2.4. *Let $h : (F_\bullet, \partial_\bullet^F) \rightarrow (G_\bullet, \partial_\bullet^G)$ be a map of complexes. For all n , h_n restricts to homomorphisms $B_n(h) : B_n(F_\bullet) \rightarrow B_n(G_\bullet)$ and $Z_n(h) : Z_n(F_\bullet) \rightarrow Z_n(G_\bullet)$. As a consequence, h induces homomorphisms on homology $H_n(h) : H_n(F_\bullet) \rightarrow H_n(G_\bullet)$.*

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

$$h_n(a) = h_n \partial_{n+1}^F(b) = \partial_{n+1}^G h_{n+1}(b) \in \text{im } \partial_{n+1}^G = B_n(G_\bullet).$$

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

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

so $h_n(a) \in \ker \partial_n^G = Z_n(G_\bullet)$. Finally, the restriction of h_n to $Z_n(F_\bullet) \rightarrow Z_n(G_\bullet)$ sends $B_n(F_\bullet)$ into $B_n(G_\bullet)$, and thus it induces a well-defined homomorphism on the quotients $H_n(F_\bullet) \rightarrow H_n(G_\bullet)$. \square

Definition 2.5. Let $h : (F_\bullet, \partial_\bullet^F) \rightarrow (G_\bullet, \partial_\bullet^G)$ be a map of complexes. We call the map

$$\begin{aligned} H_n(h) : H_n(F_\bullet) &\longrightarrow H_n(G_\bullet) \\ a + B_n(F) &\mapsto h_n(a) + B_n(G) \end{aligned}$$

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 n th homology is a functor

$$H_n : \text{Ch}(R) \rightarrow R\text{-Mod}$$

which takes each map of complexes $h : F_\bullet \rightarrow G_\bullet$ to the R -module homomorphism

$$H_n(h) : H_n(F_\bullet) \rightarrow H_n(G_\bullet).$$

Definition 2.6. A map of chain complexes h is a **quasi-isomorphism** if it induces an isomorphism in homology, meaning $H_n(h)$ 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 2.7. Note that saying that if f is a quasi-isomorphism between F and G is a stronger statement than the fact that $H_n(F) \cong H_n(G)$ for all n : it also says that there are such isomorphisms that are all induced by f .

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

Exercise 25. Let π denote the projection map from \mathbb{Z} to $\mathbb{Z}/2\mathbb{Z}$. The chain map

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

is a quasi-isomorphism.

Definition 2.8. Let $f, g: F \rightarrow G$ be maps of complexes. A **homotopy**, sometimes referred to as a **chain homotopy**, between f and g is a sequence of maps $h_n: F_n \rightarrow G_{n+1}$

$$\begin{array}{ccccccc} \cdots & \xrightarrow{\partial_{n+2}} & F_{n+1} & \xrightarrow{\partial_{n+1}} & F_n & \xrightarrow{\partial_n} & F_{n-1} & \xrightarrow{\partial_{n-1}} & \cdots \\ & & \downarrow g_{n+1} & \searrow h_n & \downarrow g_n & \searrow h_{n-1} & \downarrow g_{n-1} & & \\ \cdots & \xrightarrow{\partial_{n+2}} & G_{n+1} & \xrightarrow{\partial_{n+1}} & G_n & \xrightarrow{\partial_n} & G_{n-1} & \xrightarrow{\partial_{n-1}} & \cdots \end{array}$$

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 26. 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 2.9. Let $f, g: (F_\bullet, \partial_\bullet^F) \rightarrow (G_\bullet, \partial_\bullet^G)$ be maps of complexes. If f is homotopic to g , then $H_n(f) = H_n(g)$ for all n . In particular, every nullhomotopic map induces the zero map in homology.

Proof. Let $f, g: (F_\bullet, \partial_\bullet^F) \rightarrow (G_\bullet, \partial_\bullet^G)$ 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

$$(f - g)_n(a) = \partial_{n+1}h_n + \underbrace{h_{n-1}\partial_n(a)}_0 = \partial_{n+1}(h_n(a)) \in B_n(G_\bullet).$$

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 $H_n(f + h) = H_n(f) + H_n(h)$, which we leave as an exercise to be further explored in ??.

Notice, however, that the converse is fall: 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 27. Consider the following map of complexes:

$$\begin{array}{ccccccc} \cdots & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & \mathbb{Z}/2 & \longrightarrow & 0 & \longrightarrow & \cdots \\ & & \downarrow & & \downarrow & & \downarrow 2 & & \downarrow & & \\ \cdots & \longrightarrow & 0 & \longrightarrow & \mathbb{Z}/4\mathbb{Z} & \xrightarrow{2} & \mathbb{Z}/4\mathbb{Z} & \longrightarrow & 0 & \longrightarrow & \cdots \end{array}$$

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

Definition 2.10. If $f : (F_\bullet, \partial_\bullet^F) \longrightarrow (G_\bullet, \partial_\bullet^G)$ and $g : (G_\bullet, \partial_\bullet^G) \longrightarrow (F_\bullet, \partial_\bullet^F)$ are maps of complexes such that fg is homotopic to the identity map on $(G_\bullet, \partial_\bullet^G)$ and gf is homotopic to the identity chain map on $(F_\bullet, \partial_\bullet^F)$, we say that f and g are **homotopy equivalences** and $(F_\bullet, \partial_\bullet^F)$ and $(G_\bullet, \partial_\bullet^G)$ are **homotopy equivalent**.

Corollary 2.11. *Homotopy equivalences are quasi-isomorphisms.*

Proof. If $f : (F_\bullet, \partial_\bullet^F) \longrightarrow (G_\bullet, \partial_\bullet^G)$ and $g : (G_\bullet, \partial_\bullet^G) \longrightarrow (F_\bullet, \partial_\bullet^F)$ are such that fg is homotopic to 1_{G_\bullet} and gf is homotopic to 1_{F_\bullet} , then by Lemma 2.9 the map fg induces the identity map on homology. So for all n we have

$$H_n(f) H_n(g) = H_n(fg) = H_n(1) = 1.$$

Therefore, $H_n(f)$ and $H_n(g)$ must both be isomorphisms. \square

The converse is false.

Exercise 28. Let π denote the projection map from \mathbb{Z} to $\mathbb{Z}/2\mathbb{Z}$. The chain map

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

is a quasi-isomorphism but not a homotopy equivalence.

Remark 2.12. The relation $F \simeq G$, meaning “there is a quasi-isomorphism from F to G ”, is not symmetric: in Exercise 28, 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 and exact sequences of complexes. While we will later formalize this a little better when we discover that $\text{Ch}(R)$ is an abelian category, let’s for now give quick definitions that we can use.

Definition 2.13. 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 $\iota_n : B_n \subseteq C_n$ 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 29. 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 $C_n/B_n \rightarrow C_{n-1}/B_{n-1}$ 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 2.14. Given any map of complexes $f: B_\bullet \rightarrow C_\bullet$, 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 30. Show that the kernel, image, and cokernel of a complex map are indeed complexes.

Definition 2.15. A **complex** in $\operatorname{Ch}(R)$ is a sequence of complexes of R -modules C^n and chain maps $d_n: C^n \rightarrow C^{n-1}$ 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

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

where the n th column corresponds to the complex C^n , and every row is also a complex.

Given a complex C in $\operatorname{Ch}(R)$, 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 2.16. A **short exact sequence** of complexes is an exact complex in $\text{Ch}(R)$ 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

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

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

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

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 2.18. 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

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

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

Lemma 2.19 (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 $q: B \longrightarrow A$ such that $qf = \text{id}_A$.*
- b) *There exists a homomorphism of R -modules $r: C \longrightarrow B$ such that $gr = \text{id}_C$.*
- c) *The short exact sequence splits.*

Definition 2.20. 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**.

Remark 2.21. In the split short exact sequence

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

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

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

$$\begin{array}{ccccccc} 0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C \longrightarrow 0 \\ & & \downarrow a & & \downarrow b & & \downarrow c \\ 0 & \longrightarrow & A & \xrightarrow{i} & A \oplus C & \xrightarrow{p} & C \longrightarrow 0, \\ & & & \swarrow \pi & \nwarrow j & & \end{array}$$

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 $j: C \longrightarrow A \oplus C$ be the inclusion onto the first component. Now consider the maps $q := a^{-1}\pi b$ and $r := b^{-1}jc$. Then

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

and

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

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

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

so $b - fq(b) \in \ker q$. This shows that $B = \text{im } f + \ker q$. Moreover, if $f(a) \in \ker q$, then $a = qf(a) = 0$, so $\text{im } f \cap \ker q = 0$, and $B = \text{im } f \oplus \ker q$. 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 $\text{im } f \cong A$, so we conclude that $B \cong A \oplus C$, via the isomorphism φ given by

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

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

$$\begin{array}{ccccccc} 0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C \longrightarrow 0 \\ & & \parallel & & \downarrow \varphi & & \parallel \\ 0 & \longrightarrow & A & \xrightarrow{i} & A \oplus C & \xrightarrow{p} & C \longrightarrow 0. \end{array}$$

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

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

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

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

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

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

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

$$g\psi(a, c) = \underbrace{gf(a)}_0 + \underbrace{gr(c)}_{\operatorname{id}_C} = c,$$

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

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

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

Exercise 31. 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 32. Prove the Rank-Nulity Theorem using Exercise 31: show that given any linear transformation $T: V \rightarrow W$ 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 2.22. 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 $f: \mathbb{Z} \rightarrow \mathbb{Z}$ 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 2.23. 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 $\mathbb{Z}/(4) \not\cong \mathbb{Z}/(2) \oplus \mathbb{Z}/(2)$. 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 Lemma 2.19, 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 33 (The Five Lemma). Consider the following commutative diagram of R -modules with exact rows:

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

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

Remark 2.24. Given a short exact sequence, suppose we have R -module homomorphisms q and r

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

$$\quad \quad \quad \nwarrow \quad \quad \quad \nwarrow$$

$$\quad \quad \quad q \quad \quad \quad r$$

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

$$B \xrightarrow{\varphi} A \oplus C$$

$$b \longmapsto (q(b), g(b))$$

such that the diagram

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

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 $R\text{-Mod}$ behaves better than the category of groups, and this is one of them.

Remark 2.25. The [Splitting Lemma](#) does not hold if we replace R -modules with the category \mathbf{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

$$\text{sign}: S_3 \longrightarrow \mathbb{Z}/2,$$

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

Funny enough, there is no splitting for the inclusion $A_3 \subseteq S_3$, since there are no nontrivial homomorphisms $S_3 \longrightarrow A_3$: A_3 has no elements of order 2, so a group homomorphism $S_3 \longrightarrow A_3$ must send every 2-cycle in S_3 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 34. Fix a ring R . Show that if F is a free R -module, then every short exact sequence of R -modules

$$0 \longrightarrow A \longrightarrow B \longrightarrow R \longrightarrow 0$$

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

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

If the rows of the diagram are exact, then there exists an exact sequence

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

Given $c' \in \ker h$, pick $b' \in B'$ such that $p'(b') = c'$, and $a \in A$ such that $i(a) = g(b')$. Then

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

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

$$\begin{array}{ccccccc} \ker f & \longrightarrow & \ker g & \longrightarrow & \ker h & & \\ \downarrow & & \downarrow & & \downarrow & & \\ A' & \longrightarrow & B' & \longrightarrow & C' & & \\ \downarrow f & & \downarrow g & & \downarrow h & & \\ A & \longrightarrow & B & \longrightarrow & C & & \\ \downarrow & & \downarrow & & \downarrow & & \\ \operatorname{coker} f & \longrightarrow & \operatorname{coker} g & \longrightarrow & \operatorname{coker} h & & \end{array}$$

∂

Definition 2.27. The map ∂ in the Snake Lemma is the **connecting homomorphism**.

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

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

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

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

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 $a' \in \ker f$. An 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

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

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

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

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

$$\begin{array}{ccccccc}
 & & & & c' \in \ker h & & \\
 & & & & \downarrow & & \\
 A' & \xrightarrow{i'} & b' \in B' & \xrightarrow{p'} & c' \in C' & & \\
 \downarrow f & & \downarrow g & & \downarrow h & & \\
 a \in A & \xrightarrow{i} & g(b') \in B & \xrightarrow{p} & 0 \in C & \implies & g(b') \in \ker p = \operatorname{im} i \\
 \downarrow & & & & & & \\
 a + \operatorname{im} f \in \operatorname{coker} f & & & & & &
 \end{array}$$

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 δ 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

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

Therefore,

$$\delta(c'_1) + \delta(c'_2) = (a_1 + \operatorname{im} f) + (a_2 + \operatorname{im} f) = (a_1 + a_2) + \operatorname{im} f = \delta(c'_1 + c'_2).$$

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 $\delta(0)$ is well-defined. Ultimately, our definition of δ only involves one choice, when we pick $b' \in B'$ such that $p'(b') = 0$; we need to show that $\delta(0)$ 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

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

so $\delta(0)$ is defined as $a + \operatorname{im} f \in \operatorname{coker} f$. Since $a = f(a') \in \operatorname{im} f$, we conclude that $a + \operatorname{im} f = 0$, so $\delta(0) = 0$ for any choice of b' .

Now consider any $c' \in \ker h$. Again, to show δ 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

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

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

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

Thus

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

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

Now we show 1): that $p'(\ker g) = \ker \partial$.

If $b' \in \ker g$, then $g(b') = 0$ and the only $a \in A$ such that $i(a) = g(b') = 0$ is $a = 0$. Therefore, $\delta(p'(b')) = 0$, so $p'(\ker g) \subseteq \ker \delta$. 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 $p'i' = 0$, $p'(b' - i'(a')) = p'(b') = c'$, so $c' \in \operatorname{im} p'$. 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 $a \in A$. The statement $i(a + \operatorname{im} f) = 0$ lifts to B as $i(a) \in \operatorname{im} g$, so we can choose $b' \in B'$ such that $g(b') = i(a)$. Then $\partial(p'(b')) = a + \operatorname{im} f$, and moreover $p'(b')$ is in $\ker h$, since by commutativity we have $hp'(b') = pg(b') = pi(b') = 0$. This shows that $\ker(\operatorname{coker} f \xrightarrow{i} \operatorname{coker} g) \subseteq \operatorname{im} \partial$. Finally, if c', b' , and a are as in the diagram above,

$$i(a + \operatorname{im} f) = g(b') + \operatorname{im} g = 0,$$

so

$$\operatorname{im} \partial \subseteq \ker(\operatorname{coker} f \xrightarrow{i} \operatorname{coker} g).$$

□

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 2.28 (Long exact sequence in homology). *Given a short exact sequence in $\text{Ch}(R)$*

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

there are connecting homomorphisms $\partial : H_n(C) \longrightarrow H_{n-1}(A)$ such that

$$\cdots \longrightarrow H_{n+1}(C) \xrightarrow{\partial} H_n(A) \xrightarrow{f} H_n(B) \xrightarrow{g} H_n(C) \xrightarrow{\partial} H_{n-1}(A) \longrightarrow \cdots$$

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 Lemma 2.4, that f and g both take boundaries to boundaries, so that we get exact sequences

$$A_n / \text{im } d_{n+1}^A \longrightarrow B_n / \text{im } d_{n+1}^B \longrightarrow C_n / \text{im } d_{n+1}^C \longrightarrow 0.$$

Again by Lemma 2.4, the condition that f and g are maps of complexes also implies that f and g take cycles to cycles, so we get exact sequences

$$0 \longrightarrow Z_n(A) \longrightarrow Z_n(B) \longrightarrow Z_n(C).$$

Let F be one of A , B , or C . The boundary maps on F induce maps $F_n \longrightarrow Z_{n-1}(F)$ that send $\text{im } d_{n+1}$ to 0, so we get induced maps $F_n / \text{im } d_{n+1} \longrightarrow Z_{n-1}(F)$. Putting all this together, we have a commutative diagram with exact rows

$$\begin{array}{ccccccc} A_n / \text{im } d_{n+1}^A & \longrightarrow & B_n / \text{im } d_{n+1}^B & \longrightarrow & C_n / \text{im } d_{n+1}^C & \longrightarrow & 0 \\ d_n^A \downarrow & & d_n^B \downarrow & & d_n^C \downarrow & & \\ 0 \longrightarrow & Z_n(A) & \longrightarrow & Z_n(B) & \longrightarrow & Z_n(C) & \end{array}$$

For each $F = A, B, C$, the kernel of $F_n / \text{im } d_{n+1}^F \xrightarrow{d_n^F} Z_{n-1}(F)$ is $H_n(F)$, and its cokernel is $Z_{n-1}(F) / \text{im } d_n^F = H_{n-1}(F)$. The Snake Lemma now gives us exact sequences

$$H_n(A) \longrightarrow H_n(B) \longrightarrow H_n(C) \xrightarrow{\partial} H_{n-1}(A) \longrightarrow H_{n-1}(B) \longrightarrow H_{n-1}(C)$$

Finally, we glue all these together to obtain the long exact sequence in homology. □

Remark 2.29. It's helpful to carefully consider how we compute the connecting homomorphisms in the long exact sequence, which we can easily put together from the proof of the Snake Lemma. Suppose that $c \in \ker d_{n+1}^C$. 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 $d_{n+1}^B(b) \in \ker g_n$, we can find $a \in A_n$ with $f_n(a) = d_{n+1}^B(b)$. Finally, $\partial(c) = a + \text{im } d_{n+1}^A$.

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 2.30. 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

$$0 \longrightarrow \ker f_n \xrightarrow{i} C_n \xrightarrow{\pi} \text{coker } f_{n+1} \longrightarrow 0.$$

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 $\text{im } i = \ker f_n = \text{im } f_{n+1} = \ker \pi$.

The long exact sequence in homology is natural.

Theorem 2.31 (Naturality of the long exact sequence in homology). *Any commutative diagram in $\text{Ch}(R)$*

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

with exact rows induces a commutative diagram where the rows are long exact sequences

$$\begin{array}{ccccccccccc} \cdots & \longrightarrow & H_{n+1}(C) & \xrightarrow{\partial} & H_n(A) & \xrightarrow{i} & H_n(B) & \xrightarrow{p} & H_n(C) & \xrightarrow{\partial} & H_{n-1}(A) & \longrightarrow & \cdots \\ & & h \downarrow & & f \downarrow & & g \downarrow & & h \downarrow & & f \downarrow & & \\ \cdots & \longrightarrow & H_{n+1}(C') & \xrightarrow{\partial'} & H_n(A') & \xrightarrow{i'} & H_n(B) & \xrightarrow{p'} & H_n(C) & \xrightarrow{\partial'} & H_{n-1}(A) & \longrightarrow & \cdots \end{array}$$

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 2.28. 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

$$\begin{array}{ccccc} H_n(A) & \xrightarrow{i} & H_n(B) & \xrightarrow{p} & H_n(C) \\ f \downarrow & & g \downarrow & & h \downarrow \\ H_n(A') & \xrightarrow{i'} & H_n(B) & \xrightarrow{p'} & H_n(C) \end{array}$$

follows from the fact that H_n is a functor, so we only need to check commutativity of the square

$$\begin{array}{ccc} H_n(C) & \xrightarrow{\partial} & H_{n-1}(A) \\ h \downarrow & & \downarrow f \\ H_n(C) & \xrightarrow{\partial'} & H_{n-1}(A) \end{array}$$

that involves the connecting homomorphisms ∂ and ∂' . Consider the following commutative diagram:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A_n & \xrightarrow{i} & B_n & \xrightarrow{p} & C_n & \longrightarrow & 0 \\ & & \searrow d & & \searrow d & & \searrow d & & \\ 0 & \longrightarrow & A_{n-1} & \longrightarrow & B_{n-1} & \longrightarrow & C_{n-1} & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ & & A'_n & \longrightarrow & B'_n & \longrightarrow & C'_n & \longrightarrow & 0 \\ & & \searrow d & & \searrow d & & \searrow d & & \\ 0 & \longrightarrow & A'_{n-1} & \xrightarrow{i'} & B'_{n-1} & \xrightarrow{p'} & C'_{n-1} & \longrightarrow & 0 \end{array}$$

Given $c \in \ker(d_n: C_n \rightarrow C_{n-1})$, we need to check that $f_{n-1}(\partial(c)) = \partial'h_n(c)$ in $H_{n-1}(A)$. 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 $\partial(c) = a + \text{im } d_n \in H_{n-1}(A)$. So $f_{n-1}\partial(c) = f_{n-1}(a) + \text{im } d_n$. 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 top square

$$\begin{array}{ccc} B_n & \xrightarrow{p_n} & C_n \\ g_n \downarrow & & \downarrow h_n \\ B'_n & \xrightarrow{p'_n} & C'_n \end{array}$$

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 $\partial'(h(c)) = a' + \text{im } d_n \in H_{n-1}(A')$. By commutativity of the middle square

$$\begin{array}{ccc} B_n & \xrightarrow{d_n} & B_{n-1} \\ g_n \downarrow & & \downarrow g_{n-1} \\ B'_n & \xrightarrow{d_n} & B'_{n-1} \end{array}$$

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

$$\begin{array}{ccc} A_{n-1} & \xrightarrow{i_{n-1}} & B_{n-1} \\ f_{n-1} \downarrow & & \downarrow g_{n-1} \\ A'_{n-1} & \xrightarrow{i'_{n-1}} & B'_{n-1} \end{array}$$

we have

$$i'_{n-1}f_{n-1}(a) = g_{n-1}i_{n-1}(a) = d_n(b').$$

So we can take $a' = f_{n-1}(a)$. Finally, this means $\partial'(h_n(c)) = f_{n-1}(a) + \text{im } d_{n-1}$, as we wanted to prove. \square

Remark 2.32. Let

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

be a short exact sequence in $\text{Ch}(R)$. We can think of Theorem 2.31 as saying that the induced maps on homology $i_*: H_n(A) \rightarrow H_n(B)$ and $p_*: H_n(B) \rightarrow H_n(C)$ and the connecting homomorphism $\partial: H_n(C) \rightarrow H_{n-1}(A)$ are all natural transformations. More precisely, consider the category **SES** of short exact sequences of R -modules, which is a full subcategory of $\text{Ch}(R)$. Homology gives us functors **SES** \rightarrow $R\text{-Mod}$ that given a short exact sequence

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

return the R -modules $H_n(A)$, $H_n(B)$, or $H_n(C)$. A map between two short exact sequences then induces R -module homomorphisms between the corresponding homologies. With this framework, Theorem 2.31 says that $i_*: H_n(A) \rightarrow H_n(B)$, and $p_*: H_n(B) \rightarrow H_n(C)$ and the connecting homomorphism $\partial: H_n(C) \rightarrow H_{n-1}(A)$ are all natural transformations between the corresponding homology functors.

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 if 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 $b = 0$.

For some familiar examples, $M_n(R)$ (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 $f: R \rightarrow S$ 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 $f: \mathbb{R} \rightarrow M_2(\mathbb{R})$ 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 35. For any ring R , there exists a unique homomorphism $\mathbb{Z} \rightarrow R$.

Definition A.6. A subset S of a ring R is a **subring** of R if it is a ring under the same addition and multiplication operations and $1_R = 1_S$.

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

Definition A.7. Let R be a ring. A subset I of R is an **ideal** if:

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

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

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

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

Remark A.8. Every ring R has two **trivial ideals**: R itself and the zero ideal $(0) = \{0\}$.

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

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.9. Let R be a ring with $1 \neq 0$. A **left R -module** is an abelian group $(M, +)$ together with an action $R \times M \rightarrow M$ 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 $M \times R \rightarrow M$, $(m, r) \mapsto mr$, 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.10. If R is a commutative ring, then any left R -module M may be regarded as a right R -module by setting $mr := rm$. Likewise, any right R -module may be regarded as a left R -module. Thus for commutative rings, we just refer to modules, and not left or right modules.

The definitions of submodule, quotient of modules, and homomorphism of modules are very natural and easy to guess, but here they are.

Definition A.11. 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** $M \rightarrow M/N$ induced by sending 1 to its equivalence class.

Example A.12. 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.13. The \mathbb{Z} -modules are precisely all the abelian groups.

Example A.14. When we think of the ring R as a module over itself, the submodules of R are precisely the ideals of R .

Theorem A.15 (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.16. 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.17. Every vector space is a free module.

Remark A.18. 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

$$\begin{aligned} \bigoplus_{i \in I} R &\xrightarrow{\pi} M \\ (r_i)_{i \in I} &\longrightarrow \sum_i r_i \gamma_i \end{aligned}$$

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** $\bigoplus_{i \in I} R \xrightarrow{\pi} M$ 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.19. Given a set of generators for an R -module M , any homomorphism of R -modules $M \rightarrow N$ 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.20. 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

- $A \hookrightarrow B$, 2
- $A \twoheadrightarrow B$, 2
- $A_1 \amalg \cdots \amalg A_n$, 25
- $A_1 \times \cdots \times A_n$, 25
- $B_n(C_\bullet)$, 3
- $C \simeq D$, 47
- R -module, 67
- $Z_n(C_\bullet)$, 3
- $\text{Ch}(\mathbf{Ab})$, 46
- $H_i(C_\bullet)$, 3
- colim, 38
- $\coprod_i A_i$, 25
- $\text{im } f$, 50
- $\ker f$, 50
- lim, 38
- \mathcal{C}^{op} , 11
- $\mathcal{D}^{\mathcal{C}}$, 18
- $\prod_i A_i$, 25
- $\varinjlim M_i$, 34
- $\varprojlim M_i$, 30
- $f \simeq g$, 48
- f_* , 47
- n -boundary, 3
- n -cycle, 3
- $R\text{-Mod}$, 6
- $R\text{-mod}$, 6
- $\text{Mat-}k$, 8
- $\text{Vect-}k$, 6
- abelianization of a group, 13
- absorption, 66
- adjoint functors, 43
- basis, 68
- boundaries, 3
- bounded complex, 1
- category, 6
- category of chain complexes, 46
- chain complex, 1
- chain homotopy, 48
- chain map, 46
- cocone over a diagram, 37
- cohomology, 3
- $\text{coker}(f)$, 2
- cokernel, 2
- cokernel of a map of complexes, 50
- colimit, 34, 38
- commutative ring, 65
- complex, 1
- complex of R -modules, 3
- complex of complexes, 50
- concrete category, 7
- cone over a diagram, 37
- connecting homomorphism, 57
- constant functor, 13
- contravariant functor, 12
- converts colimits to limits, 39
- converts limits to colimits, 39
- coproduct, 25
- covariant functor, 12
- cycles, 3
- diagram, 37
- differentials, 1
- direct limit, 34
- direct system, 33
- discrete partial order, 29
- division ring, 65
- domain, 66
- dual notions in category theory, 11
- dual vector space, 13

embedding (category theory), 14
 empty category, 7
 epi, 9
 epimorphism, 9
 equivalence of categories, 19
 equivalent categories, 19
 essentially surjective functor, 14
 exact sequence, 2
 exact sequence of modules, 2

 faithful functor, 14
 field, 65
 finitely generated module, 68
 forgetful functor, 13
 free module, 68
 full functor, 14
 full subcategory, 10
 fully faithful functor, 14
 functor, 12
 functor category, 18
 functor represented by, 15

 generating set, 68
 generators for an R -module, 68

 Hom functors, 15
 homological degree, 1
 homology, 3
 homology functors, 47
 homomorphism of R -modules, 67
 homotopic, 48
 homotopy, 48
 homotopy classes, 48
 homotopy equivalence, 49
 homotopy equivalent, 49
 homotopy type, 48

 ideal, 66
 image of a map of complexes, 50
 index category of a diagram, 37
 induced map in homology, 47
 initial object, 10
 integral domain, 66
 inverse arrow, 9
 inverse limit, 30
 inverse system, 29

isomorphic objects, 9
 isomorphism (category theory), 9

 kernel of a map of complexes, 50

 left R -module, 67
 left adjoint functor, 43
 left ideal, 66
 left inverse, 8
 left invertible arrow, 8
 limit, 30, 38
 locally small category, 7
 long exact sequence in homology, 61

 map of R -modules, 67
 map of complexes, 46
 monic arrow, 9
 mono, 9
 monomorphism, 9

 natural isomorphism, 17
 natural transformation, 16
 nontrivial ideal, 66
 nullhomotopic, 48

 opposite category, 11

 presentation, 68
 preserves colimits, 39
 preserves limits, 39
 product, 25
 product category, 42
 projections, 25
 proper ideal, 66
 pullback, 35
 pullback diagram, 36
 pushout, 36
 pushout diagram, 36

 quasi-isomorphic complexes, 47
 quasi-isomorphism, 47
 quotient of complexes, 49
 quotient of modules, 67

 reflects and creates isos, 14
 relation, 68
 representable functor, 19

- right R -module, 67
- right adjoint functor, 43
- right ideal, 66
- right inverse, 9
- ring, 65
- ring homomorphism, 66
- semigroup, 8
- sends colimits to limits, 39
- sends limits to colimits, 39
- ses, 2
- short exact sequence, 2
- short exact sequence of complexes, 51
- small category, 7
- Snake Lemma, 57
- split short exact sequence, 51
- splittings, 52
- subcomplex, 49
- submodule, 67
- subring, 66
- terminal object, 10
- trivial ideals, 66
- trivial short exact sequences, 51
- universal arrow, 40
- universal element, 40
- universal property, 40
- Yoneda Lemma, 20
- zero object, 10
- zerodivisors, 66

Bibliography

- [DF04] David S. Dummit and Richard M. Foote. *Abstract algebra*. Wiley, 3rd ed edition, 2004.
- [Mac50] Saunders MacLane. Duality for groups. *Bulletin of the American Mathematical Society*, 56(6):485 – 516, 1950.
- [ML98] Saunders Mac Lane. *Categories for the working mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1998.
- [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.