# Category Theory

Luke Trujillo

# Contents

## Acknowledgments

I want to thank Professor Dagan Karp at Harvey Mudd College, my undergraduate adviser, for being an amazing mentor and for providing me extremely helpful advice and support during my time at Harvey Mudd. Professor Karp also supervised an independent study for which a large portion of this text originated from.

I also want to thank Professor Vin de Silva at Pomona College, my undergraduate thesis adviser. Through Professor de Silva's humility and patience I learned a great deal of category theory and many tecniques how to write mathematics clearly. I also would have never produced Section 7.4 or the chapter on Persistence Modules without his feedback.

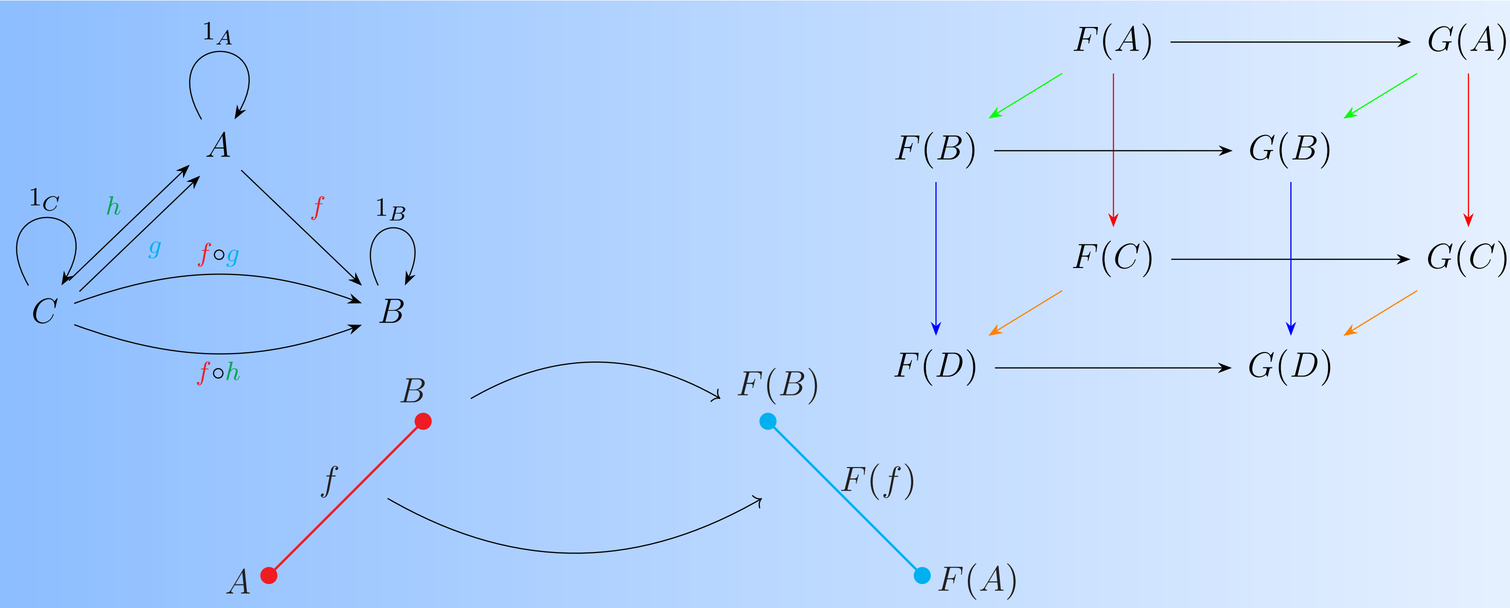
I'd also like to thank Daniel Donnelley Jr for catching dozens of typos and statements that were plain wrong; errors that could only have been caught through a thorough reading. Donnelley also offered extremely helpful feedback, and suggested many sentence revisions and section restructurings which were all in the best interest of increasing the readability of this text.

## About the Cover

The cover of the text is a long exposure shot I took from a popular hiking trail on top of Mount Baldy in Claremont, CA. You can probably tell neither the camera nor the photographer were very good, and that there's a lot of light pollution in Southern California.

# Warning!

This is a work in progress, and there are lots of typos. Some sections are being reorganized. I probably won't be "finished" for a long time. `ltrujillo@hmc.edu` .




# 1. Categories, Functors and Natural Transformations.

## 1.1 Introduction: What are the Foundations of Math?

Category theory attempts to "zoom out" of mathematical constructions and to point out the higher level relationships between different mathematical constructions. The three main concepts are categories, functors, and natural transformations, although the theory grew out of implications of these main concepts.

These main concepts were first seen in the study of algebraic topology, since it was observed that topological problems could be reduced to algebraic, and vice versa. But how? Since there was no formal notion for what it really meant to take a topological space $X$ and associate it with some group $\pi(X)$, category theory came about to formalize this.

However, as we shall soon see, category theory has a big problem. Specifically, there isn't a universally agreed upon foundation for category theory, or for mathematics in general.

**What do we mean by foundations?**

Well, consider a topological space $X$, or a group $G$, or a domain $\mathbb{R}$. Then suppose I ask you "What is $X$?" or "What is $G$" or "What is $\mathbb{R}$?" Well, you'll tell me it's a topological space, a group, or the set of real numbers and list the axioms for each object.

That is, a correct answer will characterize $X$, $G$ or $\mathbb{R}$ as a set which satisfies some axioms. But really, that's what all our mathematical objects are. So at this point, our foundations **are grounded in set theory.**

**What is set theory?**

Suppose I ask you what is set theory. While we all know there are different set theories, most people don't think about set theory axioms on a daily and won't know (like myself). But answering this question requires answering the next.

### What is a set?

We usually never have to face this question. But in developing a theory that considers relationships between different sets, we have to.

Our intuition tells us that sets $X$ are a **collection of objects, and that every collection of objects is a set**. We intuitively *think* that we can form collections of objects to create a set $X$, and that we can form intersection and unions between sets, or even compute power sets, to produce other sets. We also *think* we can also form sets such as

$$X = \{x \mid \varphi(x)\}$$

where $\varphi$ is some logical condition of inclusion. However, this leads to paradoxes, one of the most famous known as Russel's Paradox which we can describe as follows.
**Russel's Paradox.** Let $X$ be a set such that

$$X = \{A \text{ is a set } \mid A \text{ is not a member of itself.}\}$$

Now observe the following.

1. If $X \in X$, then consequently $X$ is not a member of itself. In other words, if $X \in X$, then $X \notin X$.

   Clearly, this is a contradiction. Since $X \in X$ is nonsense, $X \notin X$, right?

2. Suppose $X \notin X$. Then $X$ is not a member of itself, so $X \in X$ by the condition of member of $X$. In other words, $X \notin X \implies X \in X$.

See the problem here? **Not every collection of objects is a set.** So our previous notions of sets aren't correct.

Note that our trouble arose when we said that **a set is a collection of objects, and a collection of objects is a set.** This is because no, not every collection of objects is a set. Thus we need to go back and fix our definition of a set.

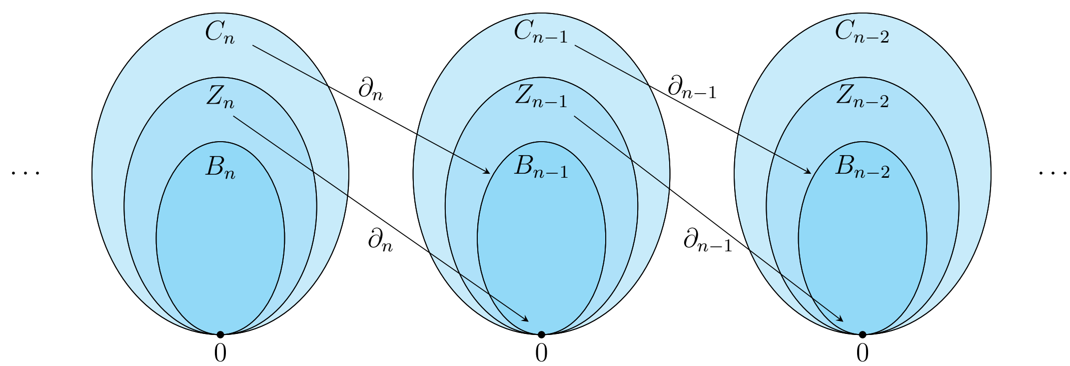
### What do we do?

This is what many mathematicians asked in the early 1900s when they identified the paradoxes that arise from our notion of a set. The result has been multiple different types of set theories, and so there isn't a clear choice on what to make our foundations. However, this isn't a huge problem for category theory. Category theory has its own core axioms, but the fact that there are different set theories simply means that such core axioms will be phrased differently under different set theories (although there are some cases where one does need to be careful with their foundations). In this text, we'll be a bit sloppy with the foundations of category theory, although we will point out where we need to be careful.

# 1.2 Motivation for Category Theory

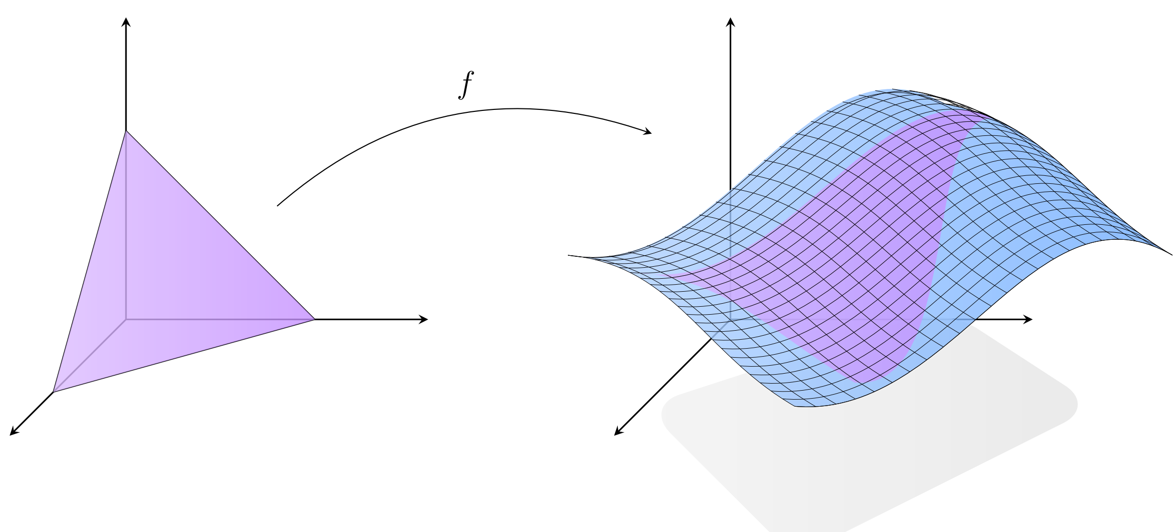
What do groups $G$, topological spaces $X$ and vector spaces $V$ have in common? *We use different letters to describe them!* Seriously, that is one major difference. Why? Because our brains are organizational and thrive off of associations, e.g., $G$ with group, $X$ with topological spaces, etc. This is great for thinking, but the mental separation of these constructions hides a bigger picture.

Let's look at what these things look like. With groups, we are often mapping between groups via group homomorphisms. For example, below we have the chain complex of abelian groups with boundary operator $\partial_n : C_n \longrightarrow C_{n-1}$, with the familiar property that $\partial_n \circ \partial_{n-1} = 0$.



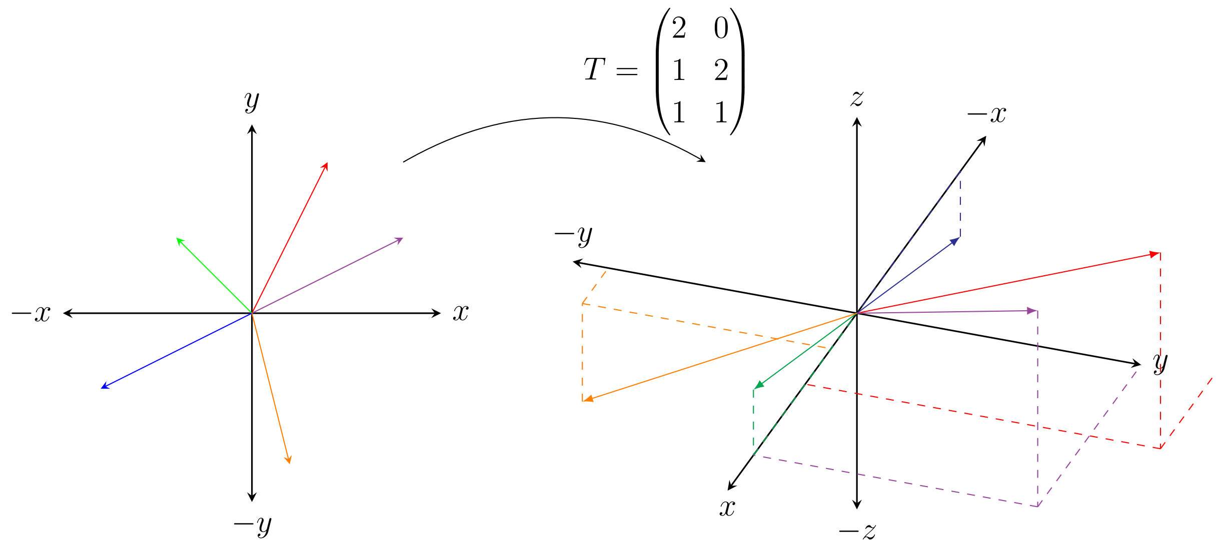

*A chain complex; the image of $\partial_n$ is $B_{n-1}$, while the kernel of $\partial_n$ is $Z_n$.*

Within topology, we are often mapping topological spaces via continuous functions.




*A 2-simplex gets embedded into a manifold in $\mathbb{R}^3$.*

With vector spaces, we often use linear transformations to map from one to another.




*Above we have $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ as a linear transformation sending the various colored vectors in $\mathbb{R}^2$ to the vectors in $\mathbb{R}^3$. The linear transformation itself is given above.*

At some point when we're learning different basic constructions in pure mathematics, we often realize that we're just repeating the same story over and over. The professor tells you about an object (usually a set) equipped with some axioms. The next thing you learn are "mappings" between such objects, which can abstractly be called *morphisms*. The characteristics of these morphism are generally the following:

**1.** There's an identity morphism.
**2.** There's a notion of composition.
**3.** Composition is associative.
**4.** Composing identities in any order with a morphism returns the same morphism.

What is it that I just described? It sounds just like a *monoid*! In the most basic sense, a monoid $M = \{x_1, x_2, \dots, \}$ is a set of elements equipped with a multiplication map

$$\cdot : M \times M \longrightarrow M \qquad (x, y) \mapsto x \cdot y$$

which is associative, and with a multiplicative identity $e$. With a monoid we see that

**1.** There's an identity $e$.
**2.** There's a notion of multiplication.
**3.** Multiplication is associative.
**4.** Multiplying $e$ in any order with an element $x$ returns $x$.

The concept of a monoid is one of the most underrated yet powerful concepts of mathematics, and for some reason it's usually ignored in algebra courses. It's an innate, fundamental *human* concept, a consequence of our physical reality. How many years have our ancestors been saying: "Let's stack stuff together and see what happens!" *Stacking three things in two different ways is the same. Stacking nothing is an "identity"*. Thus what we see is that groups, topological spaces and vector spaces are all similar in that (1) we have morphisms of interest and (2) the morphisms behave like a monoid. This notion is what category theory takes care of.

## 1.3 Category Theory Axioms.

Now we have an understanding of the fact that (1) there is no *definitive* foundation of mathematics, and therefore that (2) there is no *definitive* category theory, but rather a *definitive* set of axioms for categories. We also understand what things might look like under the axioms of category theory.

**Definition 1.3.1.** A **category** $\mathcal{C}$ consists of

- a collection of **objects** $\text{Ob}(\mathcal{C})$
- a collection of **morphisms** between objects; for any objects $A, B$, we denote the morphisms $f : A \longrightarrow B$ from $A$ to $B$ as $\text{Hom}_{\mathcal{C}}(A, B)$
- a binary operator $\circ$ known as **composition**, such that for any objects $A, B, C$,

$$\begin{aligned} \circ : \text{Hom}(A, B) \times \text{Hom}(B, C) &\longrightarrow \text{Hom}(A, C) \\ (f, g) &\mapsto (g \circ f) \end{aligned}$$

Furthermore, the following laws must be obeyed.

**(1) Identity.** For each $A \in \text{Ob}(\mathcal{C})$, there exists a distinguished morphism, called the **identity** $\text{id}_A : A \longrightarrow A$ in $\text{Hom}(\mathcal{C})$.

**(2) Closed under Composition.** If $A, B, C$ are objects, then for any $f \in \text{Hom}(A, B)$, $g \in \text{Hom}(B, C)$, there exists a morphism $h \in \text{Hom}(A, C)$ such that $h = g \circ f$.

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[rr, bend left, "h=g\circ f"] & B \arrow[r, "g"] & C \end{tikzcd}$$

**(3) Associativity under Composition.** For objects $A, B, C$ and $D$ such that

$$A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$$

we have the equality

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

**(4) Identity action.** For any $f \in \text{Hom}(\mathcal{C})$ where $f : A \longrightarrow B$ we have that

$$1_B \circ f = f = f \circ 1_A.$$

At this point, the reader is assumed to have never seen a category or has at least some vague idea. Therefore, any reasonable person would next introduce examples to clarify the above abstract nonsense. There are two types of examples we can introduce: abstract and concrete examples. We first introduce the three canonical examples, then three abstract examples. In the next section we introduce a barrage of more complicated, but *real* examples of categories in mathematics. The reader is at liberty to read the next two sections in order, in reverse, or she can skip back and forth between them.

Here we make a comment on notation. In what follows we are going to have to describe categories. To describe them, we need to tell the reader (1) what the objects are (2) what the morphisms are and (3) what composition is. Often times, (3) is implicit. Therefore our preferred format of describing an arbitrary category $\mathcal{C}$ is using a bold-faced list. An example:

> The category $\mathcal{C}$ consists of:
> **Objects.** (Here we tell you what the objects of $\mathcal{C}$ are.)
> **Morphisms.** (Here we tell you what the morphisms of $\mathcal{C}$ are.)

This is simply to avoid a lot of unnecessary words to describe a category (e.g. ”the objects of this category are... the morphisms of this category are...”).

---

**Example 1.3.2.** The canonical example of a category is the **category of sets**, denoted as **Set**, which we can describe as
**Objects.** All sets $X$.[1]
**Morphisms.** All functions between sets $f : X \longrightarrow Y$.
Because most of mathematics is based in set theory, we shall see that while this is a fairly simple category, it is one of the most useful.

---

**A tip moving forward:** When dealing with any abstract construction, it is a common strategy to keep a “canonical example” of such an abstract construction in your head. For many people, they often use **Set** as the image in their head when they imagine a category. This is fine, but one should be cautioned: in general, categorical objects are not sets. Furthemore, morphisms are in general not functions. This might be strange, but you will get used to it and it will eventually become natural to you. The moral of the story is:

> The canonical example of a category is **Set**, but *in general* the objects of an arbitrary category $\mathcal{C}$ are not sets, and the morphisms are not functions.

---

**Example 1.3.3.** The second canonical example is the **category of groups**, denoted as **Grp**. This can be described as
**Objects.** All groups $(G, \cdot)$. Here, $\cdot : G \times G \longrightarrow G$ is the group operation.
**Morphisms.** All group homomorphisms $\varphi : (G, \cdot) \longrightarrow (H, \cdot)$. Specifically, set functions $\varphi : G \longrightarrow H$ where $\varphi(g \cdot g') = \varphi(g) \cdot \varphi(g')$.
We again check this satisfies the axioms of a category.
**(1)** Each group $(G, \cdot)$ has a identity group homomorphism $\text{id}_G : (G, \cdot) \longrightarrow (G, \cdot)$ where $\text{id}_G(g) = g$.

[1] There’s a minor issue with saying this. We will address it, but not for now.

**(2)** The function composition of two group homomorphisms $\varphi : (G, \cdot) \longrightarrow (H, \cdot)$ and $\psi : (H, \cdot) \longrightarrow (K, \cdot)$ is again a group homomorphism where $(\psi \circ \varphi)(g) = \psi(\varphi(g))$. This is because

$$\begin{aligned}(\psi \circ \varphi)(g \cdot g') &= \psi(\varphi(g \cdot g'))\\ &= \psi(\varphi(g) \cdot \varphi(g'))\\ &= \psi(\varphi(g)) \cdot \psi(\varphi(g'))\\ &= (\psi \circ \varphi)(g) \cdot (\psi \circ \varphi)(g).\end{aligned}$$

**(3)** Function composition is associative; therefore, composition of group homomorphisms is associative.

**(4)** If $\varphi : (G, \cdot) \longrightarrow (H, \cdot)$ is a group homomorphism, then $\text{id}_H \circ \varphi = \varphi \circ \text{id}_G = \varphi$.

Therefore we see that this is a category. We will later see that this category possesses many convenient and interesting properties.

---

---

**Example 1.3.4.** The third canonical example is the **category of topological spaces**, denoted **Top**. We describe this as

**Objects.** All topological spaces $(X, \tau)$ where $\tau$ is a topology on the set $X$.

**Morphisms.** All continuous functions $f : (X, \tau) \longrightarrow (Y, \tau')$.

The reader can show that this too satisfies the axioms of a category.

---

We now consider some abstract examples. While abstract, they are nevertheless important examples in their own right. They also illustrate that categories can be finite, which may counter the intuition the reader might have of categories being "infinte."

---

**Example 1.3.5.** In this example we introduce the three most basic categorical structures. The first, and most important of the three, is the **single object** or **initial category 1**, which is the category where:

**Objects.** A single object, abstractly denoted as $\bullet$.

**Morphisms.** A single identity morphism $\text{id}_\bullet : \bullet \longrightarrow \bullet$.

The identity of $\bullet$ does not matter; it is an abstract object. This is similar to how a one point set is denoted as $\{*\}$ and we don't really care what $*$ is.

The second category is the **arrow category**, denoted as **2**, which we can describe as

**Objects.** Two objects $\color{blue}\bullet$ and $\color{orange}\bullet$

**Morphisms.** Two identity morphisms $\text{id}_{\color{blue}\bullet} : {\color{blue}\bullet} \longrightarrow {\color{blue}\bullet}$ and $\text{id}_{\color{orange}\bullet} : {\color{orange}\bullet} \longrightarrow {\color{orange}\bullet}$ and one nontrivial morphism $f : {\color{blue}\bullet} \longrightarrow {\color{orange}\bullet}$.

Here we color our abstract objects to clarify that these objects are distinct.

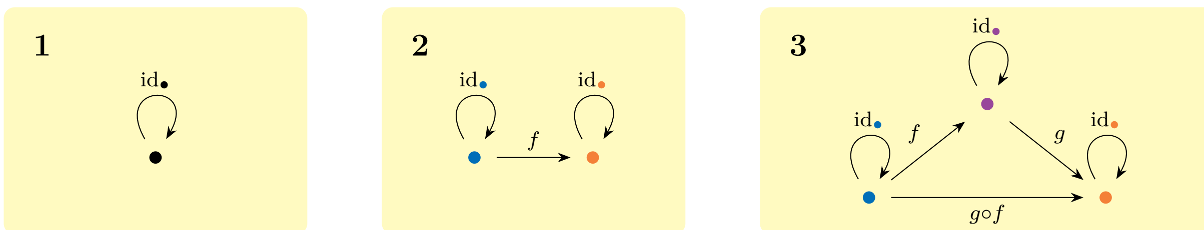
Finally, we have the category **triangle category**, denoted as **3**, which can be describe as

**Objects.** Three distinct objects $\bullet, \bullet, \bullet$

**Morphisms.** Three identity morphisms, and three nontrivial morphisms: $f : \bullet \longrightarrow \bullet$, $g : \bullet \longrightarrow \bullet$ and $h : \bullet \longrightarrow \bullet$.

In this category, we define $h = g \circ f$ so that this is closed under composition. Note that if we did not include the existence of $h$, then this would not be closed under composition, and hence it would not even be a category.

We can picture all three categories as below.




---

Our first step in category theory has been introducing the axioms and showing some simple examples. We now take our second step by moving on to more basic concepts of category theory by making a few comments about categories.

**Definition 1.3.6.** Let $\mathcal{C}$ be a category. We say that $\mathcal{C}$ is

- **Finite** if there are only finitely many objects and finitely many morphisms.
- **Locally Finite** if, for every pair of objects $A, B$, the set $\mathrm{Hom}_{\mathcal{C}}(A, B)$ is finite.
- **Small** if the collection of objects and collections of morphisms assemble into a set.
- **Locally Small** if $\mathrm{Hom}_{\mathcal{C}}(A, B)$ is a set for every pair of objects $A, B$.
- **Large** if $\mathcal{C}$ is not locally small. That is, the objects and morphisms do not form a set.

Such terminology proves to be useful, since we have seen that categories come in different sizes. For example, the categories **1**, **2**, and **3** are finite categories. However, recall Russel's Paradox, so that the collection of all sets is not a set. Therefore, **Set** is a large category.

We now introduce the concept of a *subcategory*, which is also extremely useful to include in our vocabularly.

**Definition 1.3.7.** Let $\mathcal{C}$ be a category. We say a category $\mathcal{S}$ is a **subcategory of** $\mathcal{C}$ if

**(1)** $\mathcal{S}$ is a category, with composition the same as $\mathcal{C}$

**(2)** The objects and morphisms of $\mathcal{S}$ are contained in the collection of objects and morphisms of $\mathcal{C}$.

Furthermore, we say $\mathcal{S}$ is a **full subcategory** if we additionally have that

**(3)** For each pair of objects $A, B \in \mathcal{S}$, we have that $\mathrm{Hom}_{\mathcal{S}}(A, B) = \mathrm{Hom}_{\mathcal{C}}(A, B)$.

More informally, $\mathcal{S}$ is full if it "contains all of its morphisms."

---

**Example 1.3.8.** Let **Ab** be the category described as

**Objects.** All abelian groups $(G, \cdot)$

**Morphisms.** Group homomorphisms.

Then **Ab** is a subcategory of **Grp**. Futhermore, **Ab** is a full subcategory of **Grp**. This observation also applies to

- **FinGrp**, the category of finite groups
- **FindAb**, the category finite abelian groups
- $\mathbf{Ab}_{\text{TF}}$, the category of torsion-free abelian groups

However, none of these categories are subcategories of **Set**. In fact, many categories which are based in set theory are not actually subcategories of **Set**. This is because the objects of categories such as **Grp** or **Top** are not just sets, but are sets with extra data (such as a binary operation or a topology).

---

**Example 1.3.9.** Let **Ring** be the category described as

**Objects.** Unital rings $(R, +, \cdot)$. That is, rings $R$ with a multiplicative identity 1 that is not equal to its additive identity 0.

**Morphisms.** (Unit preserving) Ring homomorphisms $\varphi : R \longrightarrow R'$. That is, functions $\varphi : R \longrightarrow R'$ such that

- $\varphi(a + b) = \varphi(a) + \varphi(b)$
- $\varphi(a \cdot b)\varphi(a) \cdot \varphi(b)$
- $\varphi(0_R) = 0_{R'}$ and $\varphi(1_R) = 1_{R'}$.

For a ring $R$ we know that $(R, +)$ is an abelian group, and we know that every ring homomorphism is technically a group homomorphism between abelian groups. However, it is not the case that **Ab** is a subcategory of **Ring**. This is because while every ring is technically an abelian group, abelian groups on their own are not rings.

---

We now introduce a convenient categorical construction which will serve to be useful to us from here on out.

**Definition 1.3.10.** Let $\mathcal{C}, \mathcal{D}$ be categories. Then we can form the **product category** where we have that

**Objects.** Pairs $(C, D)$ with $C \in \mathcal{C}$ and $D \in \mathcal{D}$.

**Morphisms.** Pairs $(f, g)$ where $f : C \longrightarrow C'$ and $g : D \longrightarrow D'$ are morphisms in $\mathcal{C}$ and $\mathcal{D}$.

To define composition in this category, suppose we have composable morphisms in $\mathcal{C}$ and $\mathcal{D}$ as below.

$$\mathcal{C} \qquad \begin{tikzcd} \cdots & C_1 \arrow[r, "f"] \arrow[rr, bend left, "f'\circ f"] & C_2 \arrow[r, "f'"] & C_3 & \cdots \end{tikzcd}$$

$$\mathcal{D} \qquad \begin{tikzcd} \cdots & D_1 \arrow[r, "g"] \arrow[rr, bend left, "g'\circ g"] & D_2 \arrow[r, "g'"] & D_3 & \cdots \end{tikzcd}$$

Then the morphisms $(f, g)$ and $(f', g')$ in $\mathcal{C} \times \mathcal{D}$ are composable too, and their composition is defined as $(f', g') \circ (f, g) = (f' \circ f, g' \circ g)$.

$$\mathcal{C} \times \mathcal{D} \qquad \begin{tikzcd} \cdots & (C_1, D_1) \arrow[r, "(f,g)"] \arrow[rr, bend left, "(f',g')\circ(f,g)=(f'\circ f,g'\circ g)"] & (C_2, D_2) \arrow[r, "(f',g')"] & (C_3, D_3) & \cdots \end{tikzcd}$$

Note that we can form even larger products of categories; we don't have to stop at two! But this will be explored later. For now, we can just be happy with this new tool because it allows us to be build new categories from the old ones that we already know.

---

**Example 1.3.11.** A useful example of a product involves the category $\mathbf{Set} \times \mathbf{Set}$ which we can describe as

**Objects.** Pairs of sets $(X, Y)$.

**Morphisms.** Pairs of functions $(f, g)$.

Such product constructions are useful because in general, algebraic operations of any kind require a product. For example, to talk about a group $(G, \cdot)$, one needs a binary operator, i.e. a function $\cdot : G \times G \longrightarrow G$. Hence to talk to generalize operations on categories, we need to talk about products. For example, with $\mathbf{Set} \times \mathbf{Set}$, we can talk about the product of two sets as a mapping $\times : \mathbf{Set} \times \mathbf{Set} \longrightarrow \mathbf{Set}$ where $(A, B) \mapsto A \times B$.

---

## 1.4 Examples of Categories

Now that we have some idea of basic categories and a few examples in mind on how they work, we introduce more examples in this section to deepen our understanding. Categories are extremely abundant in mathematics, so it is not difficult to find examples.

Without proof, we comment that the categories below truly form categories. To discuss these categories, we will use the notation in the leftmost column.

| Category | Objects | Morphisms |
|---|---|---|
| **FinSet** | Finite sets $X$ | Functions $f : X \longrightarrow Y$ |
| **Vect**$_K$ | Vector spaces over $k$ | Linear transformations $T : V \longrightarrow W$ |
| **Mon** | Monoids $(M, \cdot)$ | Monoid homomorphisms $\psi : M \longrightarrow M'$ |
| **FinGrp** | Finite Groups | Group homomorphisms $\varphi : (G, \cdot) \longrightarrow (H, \cdot)$ |
| **Ab** | Abelian Groups $(G, \cdot)$ | Group homomorphisms |
| **FinAb** | Finite Abelian Groups $(G, \cdot)$ | Group homomorphisms |
| **Ring** | Rings $(R, \cdot, +)$ | Ring homomorphisms $\varphi : (R, \cdot, +) \longrightarrow (S, \cdot, +)$ |
| **CRing** | Commutative Rings $(R, \cdot, +)$ | Ring homomorphisms |
| **Ring** | Rings $(R, \cdot, +)$ with identity $1 \neq 0$ | Ring homomorphisms |
| $R$ mod | $R$-modules $(M, +)$ | $R$-module homomorphisms |
| **Fld** | Fields $k$ | Field homomorphisms |
| **Top**$^*$ | Topological spaces $(X, x_0)$ with basepoint $x_0 \in X$ | Continuous functions preserving basepoints |
| **Toph** | Topological spaces $(X, \tau)$ | Homotopy equivalence classes |
| **Haus** | Hausdorff topological spaces $(X, \tau)$ | Continuous functions |
| **CHaus** | Compact Hausdorff topological spaces $(X, \tau)$ | Continuous functions |
| **DMan** | Differentiable manifolds $M$ | Differentiable functions $\varphi : M \longrightarrow M'$ |
| **LieAlg** | Lie algebras $\mathfrak{g}$ | Lie algebra homomorphisms |
| **Grph** | Graphs $(G, E, V)$ | Graph homomorphisms |

Now that we are aquainted with some of the categories that we'll be working with, we'll introduce more interesting categories that become useful. However, these categories are less trivial than the ones above, i.e it takes a bit of work to see how they form into categories.

---

**Example 1.4.1.** Let $X$ be a nonempty set. We can regard $X$ as a category where

**Objects.** All elements of $X$.

**Morphisms.** All morphisms are identity morphisms, and there are no morphisms between any two distinct objects.

This category, while fairly trivial, is called a **discrete category**.

---

---

**Example 1.4.2.** Consider any of the categories **Mon**, **Grp**, **Ring**, or $R$ mod . For any object of these categories, we can create the notion of a *grading*. Such a concept is a useful algebraic construction which appears in different areas of mathematics. For simplicity, we'll consider a grading on a group.

A group $G$ is said to be $\mathbb{N}$**-graded** if there exists a family of groups $G_1, G_2, \dots, G_n, \dots$ such that $G = \bigoplus_{i=1} G_i$. An example of this is the group $(\mathbb{R}[x], +)$, the single variable polynomials in one variable. To see that this is graded, observe that any polynomial $p(x)$ is of the form

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

Note that $p(x)$ consists of "components", i.e., different powers of $x$. If we let

$$\mathbb{R}_n[x] = \{ax^n \mid a \in \mathbb{R}\}$$

then we see that $\mathbb{R}[x] = \bigoplus_{i=0} \mathbb{R}_n[x]$.

More generally, if $\lambda$ is an indexing set, we say a group $G$ is $\lambda$**-graded** if there is a family of groups $G_i, i \in \lambda$ such that $G = \bigoplus_{i \in \lambda} G_i$. In addition, if $G = \bigoplus_{i \in \lambda} G_i$ and $H = \bigoplus_{i \in \lambda} H_i$ are two graded groups such that $\varphi_i : G_i \longrightarrow H_i$ is a group homomorphism, then we say $\varphi : G \longrightarrow H$ is a $\lambda$**-graded homomorphism**.

With that said, we can define the category of graded groups to be the category **GrGrp**, (read as "graded groups") described as

**Objects.** $\lambda$-graded groups $G = \bigoplus_{i \in \lambda}$ for some set $\lambda$
**Morphisms.** Graded homomorphisms between graded groups.

As we said before, this produces many graded categories, including **GrMon**, **GrRing**, $\mathbf{GrMod}_R$ etc.

---

---

**Example 1.4.3.** A monoid is a set $M$ equipped with an operation $\cdot : M \times M \longrightarrow M$ and an identity $e$ such that $e \cdot m = m \cdot e = m$ for all $m \in M$. In other words, monoids are like groups, in that we drop the requirement of an inverse.

Let $\mathcal{C}$ be a category with one object; denote this object as $\bullet$. As we have one object, we have one homset. We can then interpret $M$ as a category by setting

$$\text{Hom}_{\mathcal{C}}(\bullet, \bullet) = M.$$

Thus each $m \in M$ corresponds to a morphism. So, we can write each morphism in the category as $f_m : \bullet \longrightarrow \bullet$ for some $m \in M$. We then write $f_e = 1_\bullet$, the identity, and more generally define composition in the category as

$$f_m \circ f_{m'} = f_{m \cdot m'}.$$

Since $M$ is a monoid, and its multiplication is associative, we see that composition defined in this way is also associative. Further, for each $f_m$, we have that

$$f_e \circ f_m = f_m \circ f_e = f_m$$

since $e \cdot m = m \cdot e = m$ in the monoid $M$. Thus we can interpret monoids as one object categories.

---

**Definition 1.4.4.** A category $\mathcal{P}$ is said to be **thin** or a **preorder** if there is **at most** one morphism $f : A \longrightarrow B$ for each $A, B \in \mathcal{P}$.

The simplest thin categories are of the form below

$$\mathcal{P}$$
$$\begin{tikzcd} A \arrow[r] & B \arrow[r] & C \arrow[r] & \cdots \end{tikzcd}$$

but they may also have more complex shapes such as the category below.

$$\mathcal{P}$$
$$\begin{tikzcd}
\ddots & \vdots & \iddots \\
B & C & D \\
& A \arrow[ul] \arrow[u] \arrow[ur] \arrow[dl] \arrow[d] \arrow[dr] & \\
E & F & G \\
\iddots & \vdots & \ddots
\end{tikzcd}$$

Thin categories are very common since we often times only care about one single type of relation between any two objects. An example of such a relation is a binary relation; for any two real numbers $x, y \in \mathbb{R}$, we know that either $x \leq y$ or $y \leq x$.

This intuition is actually not very far off. Given a thin category $\mathcal{P}$, define the binary relation $\leq$ on the objects $\text{Ob}(\mathcal{P})$ as follows. For any pair of objects $A, B \in \mathcal{P}$, we have that

$$A \leq B \text{ if and only if there exists an morphism } A \longrightarrow B.$$

Some things are to be said about this relation:

- For each object $A$, there always exists a morphism $A \longrightarrow A$ (namely, the identity). This implies that $A \leq A$ for all objects $A$, so that $\leq$ is reflexive.

- If $f : A \longrightarrow B$ and $g : B \longrightarrow C$, then we have that $A \le B$ and $B \le C$. Since we may compose morphisms, we have that $g \circ f : A \longrightarrow C$. Therefore, $A \le C$, so that $\le$ $\le$ is transitive.

Hence, $\mathcal{P}$ is really just a set with a reflexive and transitive binary relation. However, this is exactly the definition of a **preorder**! Therefore, preorders $P$ can be regarded as categories with at most one morphism between any two objects, and vice versa.

Preorders can also turn into **partial orders**, which have the axiom that

$$\text{if } p \le p' \text{ and } p' \le p \text{ then } p = p'.$$

or **linear orders**, where for any $p, p'$ we have that $p \le p'$ **or** $p' \le p$.

---

**Example 1.4.5.** Here we introduce some examples of thin categories.

**Natural Numbers.** The sets $\{1, 2, \dots, n\}$ for any $n \in N$ are linear orders, each of which forms a category as pictured below.

$$\begin{tikzcd} 1 \arrow[r] \arrow[loop above] & 2 \arrow[r] \arrow[loop above] & 3 \arrow[r] \arrow[loop above] & \dots \arrow[r] & n \arrow[loop above] \end{tikzcd}$$

In this figure, the loops represent the trivial identity functions.

This example can also be generalized to include $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$.

**Subsets.** Let $X$ be a set. Then one can form a category Subsets$(X)$ where the objects are subsets of $X$ and the morphisms are inclusion morphisms. Hence, there is at most one morphism between any two sets.

Since there is at most one morphism between any two objects of the category, we see that this forms a thin category, and hence a partial ordering. What this then tells us is that subset containment determines an ordering, specifically a partial ordering.

**Open Sets.** Let $(X, \tau)$ be a topological space. Define the category $\textbf{Open}(X)$ to be the category whose objects are the open sets of $X$ and morphisms $U \longrightarrow V$ are inclusion morphisms $i : U \longrightarrow V$ whenever $U \subseteq V$. Hence, there is at most one morphism between any two open sets, so that this also forms a preorder.

**Subgroups.** Let $G$ be a group. We can similarly define the category $\textbf{SbGrp}(G)$ to be the category whose objects consists of subgroups $H \le G$, and whose morphisms are inclusion homomorphisms. This is just like the last example; and, as in the last example, there is at most one morphism between any two subgroups $H, K$ of $G$ (either $i : H \longrightarrow K$ or $i : K \longrightarrow H$). Hence, we can place a partial ordering on this, so that subgroup containment is a partial ordering.

**Ideals.** Let $R$ be a ring. Then we can form a category $\textbf{Ideals}(R)$ whose objects are the ideals $I$ of $R$ and whose morphisms are inclusion morphisms. As we've seen, this forms a thin category.

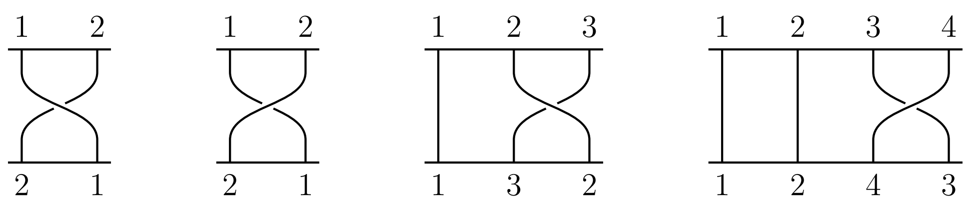
---

**Example 1.4.6.** Let $B_n$ be the set of braids on $n$ strands. Recall that $B_n$ forms a group where the group product is composition, and where the identity is simply $n$ parallel strands. Each braid group actually has a nice presentation:

$$B_n = \left\langle \sigma_1, \dots, \sigma_{n-1} \mid \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}^{(1)}, \sigma_i\sigma_j = \sigma_j\sigma_i^{(2)} \right\rangle$$

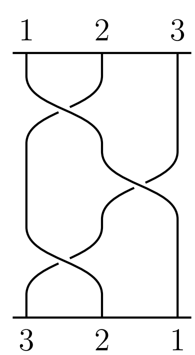
where (1) holds only when $1 \le i \le n-2$ and (2) hold only when $|i-j| > 1$. These two laws are imposed so that they match our geometric intuition, so that if we were to replace the strands with *real*, phyiscal ropes then they would behave the same way.

Each generator $\sigma_i$ is interpreted as swapping the $i$-th strand *over* the $(i+1)$-th strand, while $\sigma_i$ is swapping the $(i+1)$-th strand over the $i$-th strand. Below are some example generators.




*$\sigma_1$ on two strands; $\sigma^{-1}$ on two stands; $\sigma_2$ on three strands; $\sigma_3$ on four strands.*

The reason why we care about these generators is because every braid can be expressed by over and under crossings (although such an expression may not be unique). Now, the group multiplication in this group is simply stacking of braids. For example, the braid




can be obtained by stacking $\sigma_1, \sigma_2$ and then $\sigma_1$ again. Hence, the braid $\sigma_1\sigma_2\sigma_1$.

Now with the family of braid groups $B_1, B_2, \dots,$ we can form a category $\mathbb{B}$ as follows.

**Objects.** Positive integers $1, 2, \dots,$

**Morphisms.** For any pair of positive integers $n, m$, we have that

$$\text{Hom}_{\mathbb{B}}(n, m) = \begin{cases} B_n & \text{if } n = m \\ \varnothing & n \neq m \end{cases}$$

Hence we only have morphisms $f : n \longrightarrow m$ when $n = m$. Furthermore, each morphism is a braid. Composition is then group multiplication. The identity for each object $n$ is the identity braid of $n$ parallel strands. As group multiplication is associative, the composition in this category is associative, so we see that this truly does form a category.

---

The following examples demonstrates again that morphisms are not always functions, or mappings of some kind.

---

**Example 1.4.7.** Let $R$ be a ring with identity $1 \neq 0$. For every pair of positive integers $m, n$, let $M_{m,n}(R)$ be the set of all $m \times n$ matrices. Now recall that for an $m \times n$ matrix $A$ and a $n \times p$ matrix $B$, the product $AB$ is an $m \times p$ matrix.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{np} \end{pmatrix}$$

where $c_{ij} = \displaystyle\sum_{k=1}^{n} a_{ik}b_{kj}$. This can rephrased as saying that we have a multiplication map as below.

$$M_{m,n}(R) \times M_{n,p}(R) \longrightarrow M_{m,p}(R)$$

Since matrix multiplication is associative, we can also say that the above mapping is associative.

This however should feel sort of similar to the process of composition, say for example in **Set**, where if we have functions $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ we obtain a function $g \circ f : X \longrightarrow Z$. If we follow this intuition, we can consider a *matrix* $A$ of shape $m \times n$ as a morphism from $m \longrightarrow n$. Similarly, $B$ can be regarded a morphism from $n \longrightarrow p$. This together implies that $AB$ is a morphism from $m \longrightarrow p$. This should feel strange, because we are used to thinking of a morphism as some kind of function. But it works; we can form a category where

**Objects.** The objects are positive integers $m$.
**Morphisms.** The morphisms are matrices. Specifically, for any pair of objects $m, n$,

$$\text{Hom}_{\mathcal{C}}(m, n) = M_{m,n}(R).$$

Here, composition is simply matrix multiplication.

Observe now that our initial observation regarding matrix multiplication translates to a statement regarding whenever two matrices $A$ and $B$ are "composable" (i.e., whenever we can multiply them). That is, our mapping $M_{m,n}(R) \times M_{n,p}(R) \longrightarrow M_{m,p}$ can be rephrased as composition

$$\circ : \text{Hom}_{\mathcal{C}}(m, n) \times \text{Hom}_{\mathcal{C}}(n, p) \longrightarrow \text{Hom}_{\mathcal{C}}(m, p)$$

Associativity of matrix multiplication translates to associativity of composition. Finally, note that for each object (positive integer) $n$, the identity morphism is simply the identity matrix.

$$1_n := I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

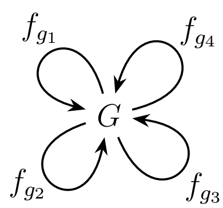
Thus we see that we have all the necessary ingredients to declare this to be a category.

---

**Example 1.4.8.** Let $G$ be a group, and recall that $G$ is equipped with some binary operator $\cdot : G \times G \longrightarrow G$ which satisfies associativity. Because this is a two-variable function on $G$ every $g \in G$ induces a map

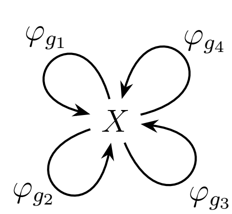
$$(-) \cdot g := f_g : G \longrightarrow G$$

This then gives rise to a collection of maps $f_g : G \longrightarrow G$ for each $g \in G$, which we can picture as below.




In particular, if $e \in G$ is the identity, then $f_e = 1_G$. Moreover, composition of these maps is associative. Thus we can think of this as a category, specifically one with one object, whose morphisms $f : G \longrightarrow G$ are induced by the elements $g \in G$. Also, note that each such map is an isomorphism, since its inverse is given by $(-) \cdot g^{-1} : G \longrightarrow G$.

Now we can step up a level of generality. Let $X$ be a set, and suppose we have a group action $\varphi : X \times G \longrightarrow X$. If we denote $\varphi(g, -) := \varphi_h : X \longrightarrow X$ for each $g \in G$, then since $\varphi$ is a group action we have that $\varphi_g \circ \varphi_{g'} = \varphi_{g \cdot g'}$ and $\varphi_e = 1_X$. Hence composition is associative and we have a well-behaved identity morphism. Usually, when we draw group actions, we think of something like this:



What we're seeing is that group actions can be phrased as a category with one object, with morphisms as isomorphisms. This generalizes our previous discussion, which makes sense since groups are trivial examples of group actions by setting $X = G$.

---

## Exercises

**1.** Let $n$ be a positive integer, and consider a group $G$ such that $g^n = 1$ for all elements $g \in G$. Show that if we take these groups to be our objects, and group homomorphisms to be our morphisms, then this forms a category $\mathbf{Grp}_n$.

**2.** Consider an infinite family of groups $G_1, G_2, \dots, G_n, \dots$ Show that we have a category $\mathbf{G}$ where

**Objects.** The positive integers $1, 2, \dots, n, \dots$
**Morphisms.** For any two positive integers $n, m$, we define

$$\text{Hom}_{\mathbf{G}}(n, m) = \begin{cases} G_n & \text{if } n = m \\ \varnothing & \text{otherwise.} \end{cases}$$

This example can be applied to many interesting families of groups, since they often come graded (i.e., they often are indexed by the positive integers.) For instance, the braid groups $B_1, B_2, \dots,$ are such an example.

**3.** Let $f : X \longrightarrow Y$ be a function between two sets. We say $f$ has the "finite-to-one" property if $f^{-1}(y)$ is always a finite set for all $y \in Y$. Show that we have a (large) category, denoted $\mathbf{Set}_{FTO}$, where

**Objects.** All sets $X$.
**Morphisms.** functions $f$ with the finite-to-one property.

**4.** Let $X$ and $Y$ be sets. A binary relation $R$ on $X$ and $Y$ is any subset of $X \times Y$. For two elements $x \in X, y \in Y$, we then write $xRy$ if $(x, y) \in R$. Binary relations can be specialized to describe functions and order relations in set theory.

Show that we can form a category where

**Objects.** All sets $X$.
**Morphisms.** For any two sets $X, Y$, we write, by abuse of notation, $R : X \longrightarrow Y$ as a morphism for each relation $R$ on $X$ and $Y$.

This category is called **Rel**, to indicate that it is the category of relations.
*Hint:* Define composition in this category as follows. Suppose $R : X \longrightarrow Y$ is a relation on $X$ and $Y$ and $P : Y \longrightarrow Z$ is a binary relation on $Y$ and $Z$. Then the composite relation

$Q : X \longrightarrow Z$ is given by

$$Q = \{(x, z) \mid \text{there exist } y \in Y \text{ such that } (x, y) \in R, (y, z) \in P\}.$$

**5.** Recall that for a two metric spaces $(M, d_M)$ and $(N, d_N)$, where $d_M : M \times M \longrightarrow M$ and $d_N : N \times N \longrightarrow N$ are the metrics, we say a function $f : M \longrightarrow N$ is a **Lipschitz-1** map with **Lipschitz constant** 1 if

$$d_N(f(x), f(y)) \leq d_M(x, y)$$

for all $x, y \in M$. Using this concept, show that we have a category where

**Objects.** Metric spaces $M$
**Morphisms.** Lipschitz-1 maps with Lipschitz constant 1.

This category is commonly denoted as **Met**.

**6.** Let $G$ be a group. We say that $G$ acts on a set $X$ if we have a function $\varphi : G \times X \longrightarrow X$ such that

- $e \cdot x = x$
- $h \cdot (g \cdot x) = (hg) \cdot x$

Such an $X$ is sometimes called a **G-set**. Note here that we represent $\varphi(g, x)$ as $g \cdot x$. Now suppose $X, Y$ are two sets for which $G$ acts on. Then we define a morphism of $G$ sets to be a function $f : X \longrightarrow Y$ such that $f(g \cdot x) = g \cdot f(x)$. Such a map is called $G$ equivariant. Show that we have a category $G$-**Sets** where

**Objects.** All $G$-sets (i.e., sets with a group action by $G$)
**Morphisms.** $G$ equivariant maps.

## 1.5 Paths and Diagrams in Categories

In this section we give an overview of the concept of a *path* and of a *diagram* within a category, which are concepts that are exactly what they sound like. This is usually a discussion that is usually glossed over, which is a huge mistake since diagrams are used everywhere in mathematics. They'll appear in nearly every section from this point on, and any good book on category theory will have dozens of diagrams. In short, they are extremely indespensible.

So, we set off to do a justice to the important concepts of paths and diagrams. However, I've kept the pragmatic reader in mind and have avoided making this discussion abstract and irrelevant.

First, we form some intuition on what exactly a diagram is. Informally, a diagram in a category $\mathcal{C}$ consists of a finite sequence of arrows between objects. Below are some diagrams.

$$
\begin{tikzcd}
A \arrow[r, "\varphi"] \arrow[d, "\sigma"'] & B \arrow[d, "\psi"] \\
C \arrow[r, "\tau"'] & D
\end{tikzcd}
\qquad
\begin{tikzcd}
& Y \arrow[dr, "g"] & \\
X \arrow[ur, "f"] \arrow[rr, "g\circ f"'] & & Z
\end{tikzcd}
$$

We can also have more complicated diagrams such as the diagrams below.

$$
\begin{tikzcd}
& & & V_4 \arrow[dr] & & & \\
& & V_5 \arrow[ur] \arrow[dl] \arrow[ddr] & & V_3 \arrow[ddl] & & \\
& V_6 \arrow[dl] & & & & V_2 \arrow[ul] \arrow[dr] \arrow[dll] & \\
V_7 & & & V \arrow[lll] \arrow[rrr] \arrow[ull] \arrow[uuu] & & & V_1
\end{tikzcd}
\qquad
\begin{tikzcd}
& A_1 \arrow[rr] \arrow[dl] \arrow[dd] & & B_1 \arrow[dl] \arrow[dd] \\
A_2 \arrow[rr] \arrow[dd] & & B_2 \arrow[dd] & \\
& A_3 \arrow[rr] \arrow[dl] & & B_3 \arrow[dl] \\
A_4 \arrow[rr] & & B_4 &
\end{tikzcd}
$$

Of course, a diagram does not really mean anything on its own; it is simply a graph[2]. A diagram requires the context of a category to have any meaning. Despite this, we can still abstract the core ingredients of what a diagram really is for a general category $\mathcal{C}$. To do so requires observing that in the diagrams above (which are the ones we care about), there are certain paths given by iterated composition. Thus we start at this concept and build upwards to define a diagram.

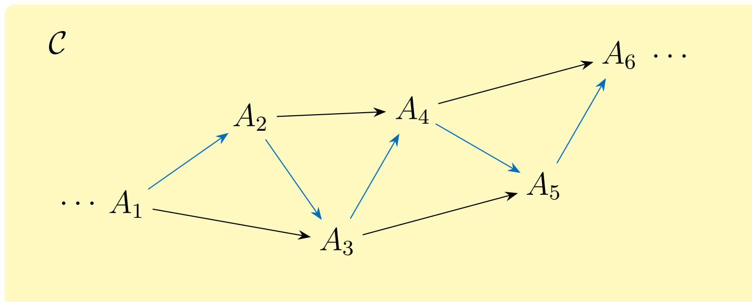
**Definition 1.5.1.** Let $\mathcal{C}$ be a category and consider two objects $A$ and $B$. A **path** $p$ in $\mathcal{C}$ of length $n$ from $A$ to $B$ consists of

- distinct objects $A_1, A_2, \dots, A_{n+1}$ with $A_1 = A$ and $A_{n+1} = B$
- a chain of morphisms $f_1 : A_1 \longrightarrow A_2, \dots, f_n : A_n \longrightarrow A_{n+1}$

[2]Technically, since a diagram can have multiple morphisms between two objects, every diagram is a "quiver." This is explored more in Chapter 2.

and we say $p = f_n \circ \cdots \circ f_1$. If two paths $p = f_n \circ \cdots f_1$ and $q = g_m \circ g_{m-1} \circ \cdots \circ g_1$ start and end at the same objects $A$ and $B$, we say $p$ and $q$ are **parallel paths**.

For example, we have a path of length five from $A_1$ to $A_6$ in some category $\mathcal{C}$ displayed below in blue.




Note that in the above picture, we will in general have many possible paths between two different objects. We now face the question: is there a way to organize this data without getting too complicated?

To answer that question, we must work with a small category in order to avoid contradictions that arise due to size issues in set theory. With that said, we propose the following definition.

**Definition 1.5.2.** Let $\mathcal{C}$ be a small category. For any two objects $A, B$, and for any positive integer $n$, define the **path set of order** $n$ from $A$ to $B$ as

$$\text{Path}^n(A, B) = \{\text{all paths } p : A \longrightarrow B \text{ of length } n\}.$$

The above definition makes sense, but admittedly it is not illuminating. Is there another perspective we can make from this?

Yes! Because paths are made of components which are inherently ordered, one way to imagine a path is as a tuple $(f_1, \dots, f_n)$ of $n$-morphisms where the codomain of $f_i$ is the domain of $f_{i+1}$. In other words, a path from $A$ to $B$ is an element of

$$\text{Hom}(A, A_1) \times \text{Hom}(A_1, A_2) \times \cdots \times \text{Hom}(A_n, B).$$

for some objects $A_1, \dots, A_n$ in $\mathcal{C}$. Therefore, we can say that

$$\text{Path}^n(A, B) = \bigcup_{A_1, \dots A_n \in \text{Ob}(\mathcal{C})} \text{Hom}(A, A_1) \times \text{Hom}(A_1, A_2) \times \cdots \times \text{Hom}(A_n, B).$$

where in the above union we vary across all objects $A_1, \dots, A_n \in \text{Ob}(\mathcal{C})$. Note that when $n = 1$, we have that $\text{Path}^n(A, B) = \text{Hom}(A, B)$. In this way, the path set can be thought of as a generalized hom-set.

**Definition 1.5.3.** A **simple diagram** $J$ in a category $\mathcal{C}$ consists of two distinguished objects $A$ and $B$, referred to as the **source** and **target** of $J$, and any finite collection of parallel paths $p_1 : A \longrightarrow B, p_2 : A \longrightarrow B, \dots, p_n : A \longrightarrow B$ of any length.

Some simple diagrams are pictured below. In the first diagram, the source and targets are $X$ and $Z$; in the second, they are $A$ and $F$; in the third, they are $V$ and $V_7$.

$$
\begin{tikzcd}
 & Y \arrow[dr] & \\
X \arrow[ur] \arrow[rr] & & Z
\end{tikzcd}
\qquad
\begin{tikzcd}
A \arrow[r] \arrow[d] & B \arrow[r] \arrow[d] & C \arrow[d] \\
D \arrow[r] & E \arrow[r] & F
\end{tikzcd}
\qquad
\begin{tikzcd}
 & & & V_4 \arrow[dl] & & & \\
 & & V_5 \arrow[dl] & & V_3 \arrow[ul] & & \\
 & V_6 \arrow[dl] & & & & V_2 \arrow[ul] & \\
V_7 & & & V \arrow[lll] \arrow[ull] \arrow[uul] \arrow[uuu] \arrow[uur] \arrow[urr] \arrow[rrr] & & & V_1 \arrow[ul]
\end{tikzcd}
$$

In many situations, simple diagrams are what we really care about. This is because often times we have two objects of interests, and we consider many possible paths between them. And in those situations, we are generally asking: are all such paths equivalent?

This is something high schoolers ask themselves all the time, and a mistake they make all the time. Let $n \geq 2$. Consider the functions

- $e : \mathbb{N} \longrightarrow \mathbb{N}$ where $f(a) = a^n$ ($e$ for exponent)
- $p : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$ where $f(a, b) = a + b$ ($p$ for plus)

Often times, they get confused and think that the paths of the diagram below are equivalent.

$$
\begin{tikzcd}
\mathbb{N} \times \mathbb{N} \arrow[r, "p"] \arrow[d, "{(e,e)}"'] & \mathbb{N} \arrow[d, "e"] \\
\mathbb{N} \times \mathbb{N} \arrow[r, "p"'] & \mathbb{N}
\end{tikzcd}
\qquad
\begin{tikzcd}
(a, b) \arrow[r, mapsto] \arrow[d, mapsto] & a + b \arrow[d, mapsto] \\
(a^n, b^n) \arrow[r, mapsto] & a^n + b^n = (a+b)^n
\end{tikzcd}
$$

Sadly, this equation does not hold generally, and the two paths of the diagram are not equivalent. Thus at this point we introduce terminology for discussing when paths are equivalent.

**Definition 1.5.4.** Let $J$ be a simple diagram in $\mathcal{C}$. If every parallel path is equal, then we say $J$ **commutes** and is a **commutative diagram**.

At this point, we should note that there is still some work to be done, since of course not all "diagrams" that we care about are simple. For example, an extremely important diagram that will eventually become engrained in your brain is pictured below on the left.[3]

$$
\begin{tikzcd}
 & Z \arrow[dl, "f"'] \arrow[d, "h"] \arrow[dr, "g"] & \\
X & X \times Y \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & Y
\end{tikzcd}
\qquad
\begin{tikzcd}
 & z \arrow[dl, mapsto] \arrow[d, mapsto] \arrow[dr, mapsto] & \\
f(z) & h(z) = (f(z), g(z)) \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & g(z)
\end{tikzcd}
$$

[3]Understanding this diagram right now is not important; there is a lot more stuff one needs to learn before we get into what this means. Long story short, it is the *universal property of a product.*

Here, the objects are sets, and the morphisms are functions; the underlying function maps are pictured above on the right.

Clearly this diagram is not simple. However, note that it is built from simple diagrams; specifically, the left and right triangles are simple diagrams. At this point, it is clear that the task of rigorously defining the notion of a diagram is reduced to defining what exactly we mean by "building" such diagrams.

## Exercises

1. Consider a category $\mathcal{C}$ with objects $A, A_0, \dots, A_n, B, B_0, B_1, \dots, B_m$. Let $A_0 = B_0 = A$ and $A_n = B_m = B$, and suppose we have a family of isomorphisms $f_i : A_{i-1} \xrightarrow{\sim} A_i$ and $g_i : B_{i-1} \xrightarrow{\sim} B_i$ as below.

$$
\begin{tikzcd}
& A_1 \arrow[r, "f_2"] & A_2 \arrow[r, "f_3"] & \cdots \arrow[dr, "f_n"] & \\
A \arrow[ur, "f_1"] \arrow[dr, "g_1"'] & & & & B \\
& B_1 \arrow[r, "g_2"'] & B_2 \arrow[r, "g_3"'] & \cdots \arrow[ur, "g_\ell"'] &
\end{tikzcd}
$$

   Suppose we have another object $C$ and isomorphisms $\varphi_i : A_i \xrightarrow{\sim} C$, $\psi_i : B_i \xrightarrow{\sim} C$ with $\psi_0 = \varphi_0$ and $\varphi_n = \psi_m$. Prove that if $\varphi_i \circ f_i = \varphi_{i+1}$ and $\psi_i \circ g_i = \psi_{i+1}$, then the above diagram is commutative in $\mathcal{C}$.

## 1.6 Functors

At this point, we really have no significant reason to care about categories. They have only so far proved to be an organizatonal tool for concepts of mathematics, but that is about it. In this section, we introduce the abstract notion of a functor which is prevalent *everywhere* in mathematics. Functors are ultimately a helpful notion which we care a lot about, but in order to define a functor we first needed to define categories. But as we have defined categories, we move on to defining functors.

**Definition 1.6.1.** Let $\mathcal{C}$ and $\mathcal{D}$ be categories. A **(covariant) functor** $F : \mathcal{C} \longrightarrow \mathcal{D}$ is a "mapping" such that

1. Every $C \in \text{Ob}(\mathcal{C})$ is assigned uniquely to some $F(C) \in \mathcal{D}$
2. Every morphism $f : C \longrightarrow C'$ in $\mathcal{C}$ is assigned uniquely to some morphism $F(f) : F(C) \longrightarrow F(C')$ in $\mathcal{D}$ such that

$$F(1_C) = 1_{F(C)} \qquad F(g \circ f) = F(g) \circ F(f)$$

If you have seen a graph homomorphism before, this definition might seem similar. This is no coincidence, and we'll see later on what the relationship between categories and graphs really are. But with that intuition in mind, we can visualize the action of a functor. Below we have arbitrary categories $\mathcal{C}$, $\mathcal{D}$, and a functor $F : \mathcal{C} \longrightarrow \mathcal{D}$.

$\mathcal{C}$

$$\begin{tikzcd}
 & & \vdots & & \\
 & \cdots & A \arrow[dl, "f"'] \arrow[dr, "g\circ f"] & \cdots & \\
\cdots & B \arrow[rr, "g"'] & & C & \cdots \\
 & \vdots & & \vdots &
\end{tikzcd}$$

$\mathcal{D}$

$$\begin{tikzcd}
 & & \vdots & & \\
 & \cdots & F(A) \arrow[dl, "F(f)"'] \arrow[dr, "F(g\circ f)"] & \cdots & \\
\cdots & F(B) \arrow[rr, "F(g)"'] & & F(C) & \cdots \\
 & \vdots & & \vdots &
\end{tikzcd}$$

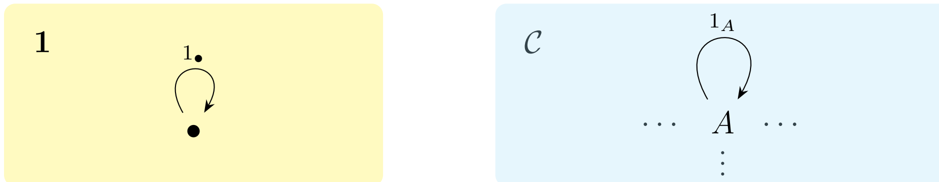
In what follows, we offer some simple and abstract examples that can get us familiar with the behavior of functors. In the next section, we do the opposite, and instead use our abstract understanding of functors to witness functors in real mathematical constructions[4].

---

**Example 1.6.2.** Denote $\mathbf{1}$ as the category with one object $\bullet$ and one identity morphism $1_\bullet : \bullet \longrightarrow \bullet$. Then for any category $\mathcal{C}$, there exists a unique functor $F : \mathcal{C} \longrightarrow \mathbf{1}$ which sends every object to $\bullet$ and every morphism to $1_\bullet$.

[4] I chose to separate this section and the next to ease the learning curve for functors; both perspectives are necessary for true understanding of a functor.

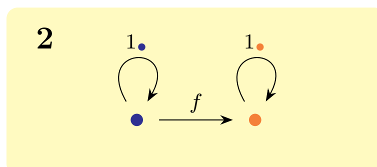
Conversely, there are many functors $F : \mathbf{1} \longrightarrow \mathcal{C}$. Since we only have $F(\bullet) = A$ for some $A \in \mathcal{C}$, and $F(1_\bullet) = 1_A$, we see that this functor simply picks out one element of $\mathcal{C}$. So these functors are in correspondence with the objects of $\mathcal{C}$; the picture below may help.



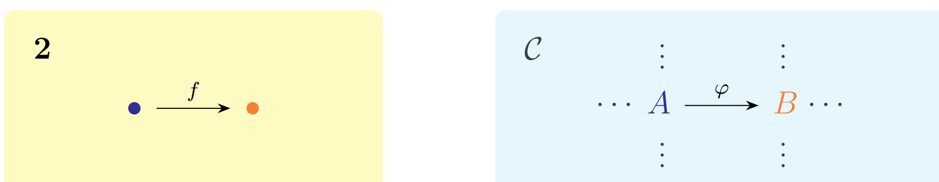

---

---

**Example 1.6.3.** Let $\mathbf{2}$ be the category with two objects $\bullet$ and $\bullet$ with one nontrivial $f : \bullet \longrightarrow \bullet$. The category can be pictured as below.




Suppose now that $\mathcal{C}$ is an arbitrary category, and that we have a functor $F : \mathbf{2} \longrightarrow \mathcal{C}$. Then note that $F(\bullet) = A$ and $F(\bullet) = B$ for some objects $A, B \in \mathcal{C}$. Hence we have that $F(f) = \varphi : A \longrightarrow B$ for some $\varphi \in \mathcal{C}$. Below we have the functor pictured.




Note we suppressed the identity morphisms. Therefore, we see that this functor simply picks out morphisms $\varphi : A \longrightarrow B$ in $\mathcal{C}$. So we can say that functors $F : \mathbf{2} \longrightarrow \mathcal{C}$ are in correspondence with the morphisms of $\mathcal{C}$.

---

Consider the very first figure of this section, Figure **??**. In that image we saw three objects $A, B, C$ get sent to $F(A), F(B), F(C)$. However, the original commutative diagram involving $f, g$ and $g \circ f$ was translated into another commutative diagram in $\mathcal{D}$ involving $F(f), F(g)$ and $F(g \circ f)$. This is because of the critical property $F(g \circ f) = F(g) \circ F(f)$ given by a functor. In fact, any commutative diagram translates to a commutative diagram under a functor.

**Proposition 1.6.4.** Let $\mathcal{C}, \mathcal{D}$ be categories with $F : \mathcal{C} \longrightarrow \mathcal{D}$ a functor. Suppose $J$ be a commutative diagram in $\mathcal{C}$. Then the diagram obtained from the image of $J$ under $F$, which we denote as $F(J)$, is commutative in $\mathcal{D}$.

*Proof.* It suffices to prove that, for any complete subdiagram $J'$ of $J$ involving any two distinct paths

$$p = f_n \circ f_{n-1} \circ \cdots \circ f_1 \qquad q = g_m \circ g_{m-1} \circ \cdots \circ g_1$$

in $J$, we have that $F(J')$ is commutative in $\mathcal{D}$. But this immediate. Since $J'$ is commutative in $\mathcal{C}$, we have that $p = q$. Hence we see that

$$F(p) = F(q) \implies F(f_n) \circ F(f_{n-1}) \cdots F(f_1) = F(g_m) \circ F(g_{m-1}) \circ \cdots \circ F(g_1).$$

by repeatedly applying the composition property of a functor. Hence $F(J')$ is commutative of $J$. Since ∎

Finally, before we move onto the next section and introduce various examples of functors across mathematics, we introduce one of the most important functors in basic category theory.

---

**Example 1.6.5.** Let $\mathcal{C}$ be a locally small category. Then for every object $A$, we obtain the **covariant hom-functor** denoted as

$$\mathrm{Hom}_{\mathcal{C}}(A, -) : \mathcal{C} \longrightarrow \mathbf{Set}.$$

where on objects $C \mapsto \mathrm{Hom}_{\mathcal{C}}(A, C)$ and on morphisms $(\varphi : C \longrightarrow C') \mapsto \varphi^* : \mathrm{Hom}_{\mathcal{C}}(A, C) \longrightarrow \mathrm{Hom}_{\mathcal{C}}(A, C'))$ where $\varphi^*$ is a function defined pointwise as

$$\varphi^*(f : A \longrightarrow C) = \varphi \circ f : A \longrightarrow C'.$$

Such a functor is naturally of interest in mathematics since it is often of interetst to consider the hom set $\mathrm{Hom}_{\mathcal{C}}(A, B)$ for some objects $A, B$ in a category $\mathcal{C}$, as it is usually the case that this set contains extra structure. For example, within topology this set is always a topological space, since families of continuous functions can be endowed with the compact open topology. In the setting of abelian groups, this set also forms an abelian group. Much of category theory can actually be done by simply "enriching" hom sets of a category with some extra structure; this is the object of enriched category theory, which we'll introduce later.

This functor in general also exhibits nice properites. For example, let $R$ be a ring. Then the sequence below

$$0 \longrightarrow M_1 \xrightarrow{\varphi} M \xrightarrow{\psi} M_2$$

is exact if and only if, for every $R$-module $N$, the sequence

$$0 \longrightarrow \mathrm{Hom}(N, M_1) \xrightarrow{\varphi^*} \mathrm{Hom}(N, M) \xrightarrow{\psi^*} \mathrm{Hom}(N, M_2)$$

is exact. This result even extends to split short exact sequences. We also have that for $R$-modules $N$, $M_1, M_2$ that

$$\mathrm{Hom}(N, M_1 \oplus M_2) \cong \mathrm{Hom}(N, M_1) \oplus \mathrm{Hom}(N, M_2).$$

This result also holds for arbitrary direct sums, so that the hom functor distributes over all direct sums. Even better, we cannot forget that the hom-functor exhibits the **tensor-hom adjunction** which states that for $R$-modules $N, M_1, M_2$

$$\mathrm{Hom}(N \otimes M_1, M_2) \cong \mathrm{Hom}(N, \mathrm{Hom}(M_1, M_2)).$$

More is to be said about this property; we'll later see that this is an example of an *adjunction*.

## 1.7 Examples and Nonexamples of Functors

Functors were not defined out of arbitrary interest. The definition of a functor was motivated by constructions that were seen in mathematics (unlike constructions in say, number theory, which are interesting in their own right). Thus in this section, we include a wide variety of different constructions in in different areas of mathematics which all fit the definition of a functor. We present examples from algebraic topology, differential geometry, topology, algebraic geometry, abstract algebra and set theory.

In short, this section is due to the fact that the only way to really understand what a functor does is to realize the definition *with examples*. It's not necessarily important to understand *all* the examples, if for instance you have never worked with differential geometry, but it would be good to get a few of them. What is more important is witnessing how such a simple definition can be so versatile and prevalent in seemingly different fields of mathematics; thus, what is important is witnessing the flexibility of functors (in addition to filling in the details of the examples and doing the exercises at the end).

### Algebraic Geometry.

---

**Example 1.7.1.** In algebraic geometry, it is often of interest to construct the **affine $n$-space** $A^n(k)$ of a field $k$. Usually, $k$ is an algebraically closed field, but it doesn't have to be.

$$A^n(k) = \{(a_0, \dots, a_{n-1}) \mid a_i \in k\}.$$

For example, when $k = \mathbb{R}$, we have that $A^n(k) = \mathbb{R}^n$. Moreover, we claim that we have a functor $A^n(-) : \mathbf{Fld} \longrightarrow \mathbf{Set}$. To see this, we need to figure out where $A^n(-)$ sends objects and morphisms.

We can first observe that $A^n(-)$ sends fields $k$ to sets $A^n(k)$. Secondly, we can observe that for a field homomorphism $\varphi : k \longrightarrow k'$, we can define the function $A^n(\varphi) : A^n(k) \longrightarrow A^n(k')$ where for each $(a_1, \dots, a_n) \in A^n(k)$ we have that

$$A^n(\varphi)(a_0, \dots, a_{n-1}) = (\varphi(a_0), \dots, \varphi(a_{n-1})).$$

The reader can show that this satisfies the rest of the axioms of a functor. Overall, we can say that we have a functor

$$A^n(-) : \mathbf{Fld} \longrightarrow \mathbf{Set}\,.$$

---

---

**Example 1.7.2.** Once the affine $n$-space is defined, the next step in algebraic geometry is to construct the **projective space** $P^n(k)$ for a field $k$. To define this, we first define an equivalence relation on $A^{n+1}(k)$. We say

$$(a_0, \dots, a_n) \sim (b_0, \dots, b_n) \text{ if there is a nonnzero } \lambda \in k \text{ such that } a_i = \lambda b_i.$$

This defines an equivalence relation on the points of $A^n(k)$. Geometrically, this equivalence relation says two points are equivalent whenever they lie on the same line passing through the origin. With this equivalence relation, we then define

$$P^n(k) = A^{n+1}/\sim = \Big\{[(a_0, \dots, a_n)] \mid (a_0, \dots, a_n) \in A^{n+1}(k)\Big\}.$$

Hence we see that $P^n(k)$ is the set of equivalence classes under this equivalence relation. Similar to the previous example, this construction is also functorial. Consider a field homomorphism $\varphi : k \longrightarrow k'$. Then we define the function $P^n(\varphi) : P^n(k) \longrightarrow P^n(k')$ where

$$P^n(\varphi)([a_0, \dots, a_n]) = [(\varphi(a_0), \dots, \varphi(a_n))].$$

However, when defining functions on a set of equivalence classes, we need to be careful. It's possible that the function could send equivalent objects to different things, so that such a fuction would not be well-defined. In this case, the above function is in fact well-defined. This is because $\varphi(\lambda a_i) = \varphi(\lambda)\varphi(a_i)$ for any $i = 0, 1, \dots, n$. Therefore we can state that we have a functor

$$P^n(-) : \mathbf{Fld} \longrightarrow \mathbf{Set}\,.$$

---

## Algebraic Topology.

---

**Example 1.7.3.** An important example of a functor arises in homology theory. For example, in singular homology theory, one considers a topological space $X$ and associates this with its $n$-th homology group.

$$X \mapsto H_n(X)$$

In a typical topology course, one then proves that if $f : X \longrightarrow Y$ is a continuous mapping between topological spaces, then $f$ induces a group homomorphism

$$H_n(f) : H_n(X) \longrightarrow H_n(Y)$$

in such a way that for a second mapping $g : Y \longrightarrow Z$, $H_n(g \circ f) = H_n(g) \circ H_n(f)$ for all $n$. Finally, we also know that $H_n(1_X) = 1_{H_n(X)}$. Therefore, what we see is that this process can be cast into the language of category theory, so that we may define a **singular homology functor**

$$H_n : \mathbf{Top} \longrightarrow \mathbf{Ab}$$

since this functorial process sends topological spaces in **Top** to abelian groups in **Ab**.

---

**Example 1.7.4.** Another example from algebraic topology can be realized from the **fundamental group**

$$\pi_1(X, x_0) = \{[x] \mid x \in X\}$$

with $x_0 \in X$, and where $[x]$ is the equivalence class of loops based at $x_0$, subject to the homotopy equivalence relation. First observe that $X \mapsto \pi_1(X)$ is in fact a mapping of objects between $\mathbf{Top}^*$ and $\mathbf{Grp}$. Second, observe that if $f : X \longrightarrow Y$ is a continuous function, then we can define a group homomorphism

$$\pi_1(f) : \pi_1(X) \longrightarrow \pi_1(Y) \qquad [x] \mapsto [f(x)].$$

Note that this is well defined since if $x \sim x'$ then there is a homotopy relation $H : X \times [0,1] \longrightarrow Y$. However, $f \circ H$ is also another homotopy relation that establishes that $f(x) \sim f(x')$; hence our group homomorphism is well defined.

Moreover, if $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ are continuous, then we can check that $\pi_1(g \circ f) = \pi_1(g) \circ \pi_1(f)$; if $[\alpha] \in \pi_1(X, x_0)$, then

$$(g \circ f)_*([\alpha]) = [(g \circ f) \circ \alpha] = [g \circ (f \circ \alpha)] = g_*([f_*([\alpha])]) = g_* \circ f_*([\alpha])$$

so that $(g \circ f)_* = g_* f_*$. Finally, we can examine how the identity map $1_X$ on a topological space acts on an element $[\alpha] \in \pi_1(X, x_0)$:

$$id_*([\alpha]) = [id \circ \alpha] = [\alpha].$$

so that it is sent to the identity homomorphism. All together, this allows us to conclude that this process is entirely functorial, so we may summarize our results by stating that

$$\pi_1 : \mathbf{Top}^* \longrightarrow \mathbf{Grp}$$

is a functor.

---

We now present two examples from differential geometry, which aren't traditionally presented as examples of functors but are nevertheless interesting in their own right.

## Differential Geometry.

---

**Example 1.7.5.** Let $M^n$ be a differentiable manifold of dimension $n$. In general, this means that there exists a family of open sets $U_\alpha \subseteq \mathbb{R}^n$ and injective mappings $\boldsymbol{x}_\alpha : U_\alpha \longrightarrow M$ for $\alpha \in \lambda$,

$\lambda$ an indexing set, with the mappings subject to various conditions[5]. Recall from differential geometry that we can associate each point $p \in M^n$ with its **tangent space** $T_p(M)$, in the following manner.

Suppose for $\alpha' \in \lambda$ we have that $\boldsymbol{x}_\alpha : U_\alpha \longrightarrow M$ is a mapping whose image contains $p$ (such an $\alpha'$ must exist). Then $T_p(M)$ has a basis

$$\left\{\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n}\right\}$$

where $\dfrac{\partial}{\partial x_i}$ is the tangent vector of the map $\boldsymbol{c}_i : U \longrightarrow M$, which simply sends $(0, \dots, 0, x_i, 0, \dots, 0)$.

Now suppose $\varphi : M_1^n \longrightarrow M_2^m$ is a differentiable mapping. Recall that the **differential** of $\varphi$ establishes a linear transformation between the vector spaces:

$$d\varphi_p : T : M_1^n \longrightarrow T_{\varphi(p)} M_2^m.$$

Consider the category $\mathbf{DMan}^*$ whose objects are pairs $(M^n, p)$ with $M^n$ a differentiable manifold and $p \in M^n$. The morphism are $(\varphi, p) : (M_1^n, p) \longrightarrow (M_2^m, q)$ with $\varphi : M_1^n \longrightarrow M_2^m$ a differentiable map and $\varphi(p) = q$. Then this process may be summarized as a functor $T_p : \mathbf{DMan}_n^* \longrightarrow \mathbf{Vect}_{\mathbb{R}}$ where

$$T : (M, p) = T_p(M)$$
$$T(\varphi : (M_1^n, p) \longrightarrow (M_2^m, \varphi(p))) = d\varphi_p : T_p(M) \longrightarrow T_{\varphi(p)} M_2^m$$

One can show that the identity map is sent to the identity linear transformation on $T_p(M)$ and that the differential respects composition, so that that the association of a manifold $M$ (with a specified point $p \in M$) to its tangent space $T_p(M)$ gives rise to a functor

$$T_p : \mathbf{DMan}^* \longrightarrow \mathbf{Vect}_{\mathbb{R}}\,.$$

---

**Example 1.7.6.** Consider again a differentiable manifold $M^n$ of dimension $n$. Recall that we may consider the **tangent bundle** $TM$ of $M$, which is the set

$$TM = \{(p, v) \mid p \in M^n \text{ and } v \in T_p(M)\}.$$

The set $TM$ simply pairs each point $p \in M^n$ with its tangent space $T_p(M)$. However, $TM$ is more than such a set; it inherits the structure of a differentiable manifold from $M$ as well. In fact, it is a manifold of dimension $2n$.

[5] There isn't a universally agreed upon set of conditions, and we won't really need to worry about them here. If the reader likes, they can consult Do Carmo's *Riemannian Geometry*, which is, and has been for a long time, the go-to differential geometry text.

Now suppose we have a differentiable mapping $\varphi : M_1^n \longrightarrow M_2^m$. Then this induces a mapping

$$(\varphi, d\varphi) : TM_1^{2n} \longrightarrow TM_2^{2m}$$
$$(\varphi, d\varphi)(p, v) = (\varphi(p), d\varphi_p(v)).$$

One can show that $(\varphi, d\varphi) : TM_1^{2n} \longrightarrow TM_2^{2m}$ is a differentiable mapping between manifolds[6] At this point we may guess that we have a functor $TB$ : **DMan** $\longrightarrow$ **DMan** ("$TB$" for "tangent bundle") where

$$TB(M^n) = TM$$
$$TB(\varphi : M_1^n \longrightarrow M_2^m) = (\varphi, d\varphi) : TM_1^{2n} \longrightarrow TM_2^{2m}.$$

To check this, we first observe that $TB(1_{M^n}) = 1_{TM^{2n}}$. Next, suppose $\varphi : M_1^n \longrightarrow M_2^m$ and $\psi : M_2^m \longrightarrow M_3^p$, and observe that

$$TB(\psi \circ \varphi) = (\psi \circ \varphi, d_{\psi\circ\varphi}) = (\psi, d_\psi) \circ (\varphi, d_\varphi) = TB(\psi) \circ TB(\varphi).$$

Note that above in the second step, we used the fact that $d_{\psi\circ\varphi} = d_\psi \circ d_\varphi$, which we know is true from the previous example. As $TB$ respects the identity and composition, we see that we do in fact have a functor

$$T : \mathbf{DMan} \longrightarrow \mathbf{DMan}$$

as desired.

---

## Topology.

---

**Example 1.7.7.** Let $X$ be a set. Recall that we can turn $X$ into a topological space $(X, \tau_d)$, where $\tau_d^{(X)}$ is the discrete topology, so that every subset of $X$ is an open set. We claim that this process is functorial, so that we have a functor

$$D : \mathbf{Set} \longrightarrow \mathbf{Top}\,.$$

This is because any function $f : X \longrightarrow Y$ extends to a continuous function $f : (X, \tau_D^{(X)}) \longrightarrow (Y, \tau_D^{(Y)})$ (hopefully the abuse of notation in $f$ is forgivable here). Hence this defines a functor, although in a simpler way than we've seen in the previous examples.

---

**Example 1.7.8.** Let $(X, \tau)$ be a topological space and consider any $x_0 \in X$. Then $(X, x_0)$

[6] I wanted to show this here, but it turned out to be just tedious definition-checking, so I don't think it's appropriate to include here (perhaps I could make/put it in an appendix...)

forms an element of **Top**$^*$. With such a space, we can consider the **loop space** of $(X, x_0)$ defined to be

$$\Omega(X) = \{\varphi : S^1 \longrightarrow X \mid \varphi \text{ is continuous and } \varphi(0) = x_0\}.$$

Here $S^1$ is the circle. As this consists of a family of continuous functions between two topological spaces, it can be endowed with the Compact Open topology to turn it into a topological space as well. Hence we claim we have a functor

$$\Omega : \mathbf{Top}^* \longrightarrow \mathbf{Top}\,.$$

To see this, one needs to first consider a morphism in **Top**$^*$, which in this case is continuous function $f : (X, x_0) \longrightarrow (Y, y_0)$ such that $f(x_0) = y_0$. This must then correspond with a continuous function $\Omega(f) : \Omega(X) \longrightarrow \Omega(Y)$. We can define this function pointwise: for each continuous $\varphi : S^1 \longrightarrow X$ such that $\varphi(0) = x_0$, we have that $\Omega(f)(\varphi) = f \circ \varphi : S_1 \longrightarrow Y$. In this case we see that $(f \circ \varphi)(0) = y_0$, and is a continuous function, so it is well-defined.

This example can be further generalized to higher loop spaces which consider continuous functions $\varphi : S^n \longrightarrow X$, rather than just having $n = 1$.

---

## Algebras, Rings, Groups.

---

**Example 1.7.9.** Recall that a **Lie Algebra** $\mathfrak{g}$ is a vector space $\mathfrak{g}$ (over a field $k$), equipped with a bilinear operation $[-,-] : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}$ such that

1. $[x, y] = -[y, x]$
2. $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0.$

Condition (2) is referred to as the **Jacobi identity**, and many familiar operations on vector spaces satisfy (1) and (2). For example, the cross product on vector spaces in $\mathbb{R}^3$ satisfy these properties.

Consider an associative algebra $A$ over a field $k$ with (associative); recall that this too has a bilinear operation $\cdot : A \times A \longrightarrow A$ with unit $e \in A$. Then we can use $A$ to create a Lie algebra $L(A)$, whose (1) underlying vector space is $A$ and (2) whose bilinear operation is $[a, b] = a \cdot b - b \cdot a$.

Now suppose $\varphi : A \longrightarrow A'$ is a morphism of algebras. Then we can associate both $A, A'$ with their Lie algebras $L(A), L(A')$. Further, we can construct a Lie Algebra morphism $L(\varphi) : L(A) \longrightarrow L(A')$, using $\varphi$, by setting $L(\varphi)(a) = \varphi(a)$. This is a morphism of Lie algebras since

$$[\varphi(a), \varphi(b)] = \varphi(a)\varphi(b) - \varphi(b)\varphi(a) = \varphi(ab - ba) = \varphi([a, b]).$$

One can then check that $L(1_A) = 1_{L(A)}$ and $L(\varphi \circ \psi) = L(\varphi) \circ L(\psi)$, so that what we have is a functor

$$L : \mathbf{Alg} \longrightarrow \mathbf{LieAlg}$$

which associates each associative algebra with its Lie algebra structure.

---

**Example 1.7.10.** Let $R$ be a commutative ring. Recall that $\text{Spec}(R)$ is the set of all prime ideals of $R$. In addition, recall that if $\varphi : R \longrightarrow S$ is a ring homomorphism and if $P$ is a prime ideal of $S$, then $\varphi^{-1}(P)$ is also a prime ideal in $R$. This then allows us to define a functor

$$\mathbf{Spec} : \mathbf{Ring} \longrightarrow \mathbf{Set}$$

where on objects $R \mapsto \text{Spec}(R)$ and on morphisms $\varphi : R \longrightarrow S \mapsto \varphi^* : \text{Spec}(S) \longrightarrow \text{Spec}(R)$ where $\varphi^*(P) = \varphi^{-1}(P)$.

However, we can go even deeper than this. Recall from algebraic geometry that $\text{Spec}(R)$ can be turned into a topological space, using the Zariski topology. However, because $\varphi^{-1}(P)$ is a prime ideal whenever $P$ is, we see that $\varphi^* : \text{Spec}(S) \longrightarrow \text{Spec}(R)$ is actually a continuous function between the topological spaces. Hence we can view this as a functor

$$\mathbf{Spec} : \mathbf{Ring} \longrightarrow \mathbf{Top}\,.$$

Usually this is phrased more naturally as a functor $\mathbf{Spec} : \mathbf{Ring} \longrightarrow \mathbf{Sch}$ where $\mathbf{Sch}$ is the category of schemes; this is simply because schemes are isomorphic to $\text{Spec}(R)$ for some $R$.

---

**Example 1.7.11.** Let $G$ be a group, and $R$ be a ring with identity. Recall from ring theory that we can form the **group ring**

$$R[G] = \left\{ \sum_{g \in G} a_g g \mid a_g \in R, \text{ all but finitely many } a_g = 0 \right\}.$$

Thus the elements are finite sums, but we have possibly infinitely many ways of adding them. Now for two elements $\alpha = \sum_{g \in G} a_k g$ and $\beta = \sum_{g \in G} b_g g$, we define ring addition and multiplication as

$$\alpha + \beta = \sum_{g \in G} (a_k + b_k) g \qquad \alpha \cdot \beta = \sum_{g \in G} \sum_{g_1 \cdot g_2 = g} (a_{g_1} b_{g_2}) g.$$

Now suppose $\varphi : G \longrightarrow H$ is any group homomorphism. With that said, we claim that $\varphi$ induces a natural ring homomorphism $\varphi^* : R[G] \longrightarrow R[H]$ between the group rings, where

$$\sum_{g \in G} a_g g \mapsto \sum_{g \in G} a_g \varphi(g).$$

Clearly this is linear and preserves scaling; less obvious is if this behaves on multiplication, although we check that below. If $\alpha, \beta$ defined as above then

$$\varphi^*(\alpha\cdot\beta) = \varphi^*\left(\sum_{g\in G}\sum_{g_1\cdot g_2=g}(a_{g_1}b_{g_2})g\right) = \sum_{g\in G}\sum_{g_1\cdot g_2=g}(a_{g_1}b_{g_2})\varphi(g) = \sum_{g\in G}a_g\varphi(g)\cdot\sum_{g\in G}b_g\varphi(g) = \varphi^*(\alpha)\cdot\varphi^*(\beta).$$

Hence we see that $\varphi^*$ is a ring homomorphism. Therefore, what we have on our hands is a functor

$$R[-] : \mathbf{Grp} \longrightarrow \mathbf{Ring}$$

Possibly, your brain may wonder: it looks like we have an assignment of rings to *functors*.

$$R \mapsto R[-] : \mathbf{Grp} \longrightarrow \mathbf{Ring}\,.$$

Perhaps this process is functorial? The answer is yes, although at the moment we don't have the necessary language to describe it; we will go over this when we introduce *functor categories*.

---

## Set Theory

---

**Example 1.7.12.** Consider the power set $\mathcal{P}(X)$ on a set $X$. Then we can create a functor $\mathcal{P} : \mathbf{Set} \longrightarrow \mathbf{Set}$ as follows.

Observe that for any set $X$, $\mathcal{P}(X)$ is of course another set. Therefore objects of **Set** are sent to **Set**, as we claim.

Now let $f : X \longrightarrow Y$ be a function between two sets $X$ and $Y$. Then we define $\mathcal{P}(f) : \mathcal{P}(X) \longrightarrow \mathcal{P}(Y)$ to be the function where

$$P(f)(S) = \{f(x) \mid x \in S\}.$$

which is clearly in $\mathcal{P}(Y)$. Now we must show that this function respects identity and composition properties.

**Identity.** Consider the identity function $\text{id}_X : X \longrightarrow X$ on a set $X$. Then observe that for any $S \in \mathcal{P}X$, we have that

$$\mathcal{P}(\text{id}_X)(S) = \{\text{id}_X(x) \mid x \in X\} = S.$$

Therefore, $\mathcal{P}(\text{id}_X) = 1_{\mathcal{P}X}$ so that $\mathcal{P}$ respects identities.

**Composition.** Let $X, Y, Z$ be sets and $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ be functions. Let $S \in \mathcal{P}(X)$.

Observe that

$$\begin{aligned}
\mathcal{P}(g \circ f)(S) &= \{(g \circ f)(x) \mid x \in S\} \\
&= \{g(f(x)) \mid x \in S\} \\
&= \{g(y) \mid y = f(x) \text{ and } x \in S\} \qquad = \mathcal{P}(g)(\{f(x) \mid x \in S\}) \\
&= \mathcal{P}(g)(\mathcal{P}(f)(S)) \\
&= (\mathcal{P}(g) \circ \mathcal{P}(f))(S).
\end{aligned}$$

Therefore we see that $\mathcal{P}(g \circ f) = \mathcal{P}(g) \circ \mathcal{P}(f)$, so that $\mathcal{P}$ describes a functor from **Set** to **Set**.

---

As we just encountered a mass of different examples of functors from different fields, one might wonder: are there other mathematical constructions which simply do not behave exactly as a functor? The answer is yes, although finding these examples is a bit tricky. The following is a well-known example, while the one after is one I haven't seen presented elsewhere.

## Non-functor Examples.

---

**Example 1.7.13.** Recall from group theory that, with every group $G$, there is an associated subgroup of $G$ called the center:

$$Z(G) = \{z \in G \mid zg = gz \text{ for all } g \in G\}.$$

By definition, $Z(G)$ is an abelian group. As every group $G$ may be associated with an abelian group $Z(G)$, one might expect that this process is functorial. One might prematurely denote this as

$$Z : \mathbf{Grp} \longrightarrow \mathbf{Ab}\,.$$

However, this is not a functor, as an issue arises with the morphisms. Consider a group homomorphism $\varphi : G \longrightarrow H$. Then for this to be a functor, we'd naturally desire a group homomorphism $Z(\varphi) : Z(G) \longrightarrow Z(H)$ between the abelian groups. The only issue is that there is no consistent way to define such a morphism from $\varphi$. The most natural way we would attempt to achieve this is by considering the restriction, but in general $\varphi\big|_{Z(G)} : G \longrightarrow H$ does not map into $Z(H)$. For example, consider the **Heisenberg Group**

$$H_3(R) = \left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \middle| a, b, c \in R \right\}$$

where $R$ is a commutative ring with identity. Observe that we can create an inclusion group homomorphism $i : H_3(R) \longrightarrow \mathrm{GL}_3(R)$. One can show that

$$Z(H_3(R)) = \left\{ \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \middle| a \in R \right\} \qquad Z(\mathrm{GL}_3(R)) = \left\{ \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix} \middle| a \in R \right\}.$$

Hence restricting the inclusion $i : H_3(R) \longrightarrow \mathrm{GL}_3(R)$ to $Z(H_3(R))$ results in a group homomorphism that does not even hit $Z(\mathrm{GL}_3(R))$ (except of course when $a = 0$). Thus there is not a general way to relate these two quantities in a functorial fashion.

---

What follows is a second example in which a process which may appear to be functorial does not turn out to be. It can, however, be adjusted to become a functor.

---

**Example 1.7.14.** Let $X$ be a set. Recall from topology that we can treat $X$ as a topological space by associating to it the **finite complement topology:**

$$\tau_{FC}^X = \{U \subseteq X \mid X - U \text{ is finite.}\}$$

With that said, one may suppose that we have a functor FinC : $\mathbf{Set} \longrightarrow \mathbf{Top}$ where $X \mapsto (X, \tau_{FC}^X)$. This would require that each function $f : X \longrightarrow Y$ extends to a continuous function $f : (X, \tau_{FC}^X) \longrightarrow (Y, \tau_{FC}^Y)$. However, for such a function to be continuous we would need that

$$\text{if } Y - V \text{ is finite then } X - f^{-1}(V) \text{ is finite.}$$

In general, this is not true. For example suppose $X$ is infinite and $Y$ is finite. Then $Y - \varnothing$ is finite, but $X - f^{-1}(\varnothing) = X$ is infinite. Hence this cannot define a functor $F : \mathbf{Set} \longrightarrow \mathbf{Top}$.

---

## Exercises

**1.** (*i.*) Let $X$ and $Y$ be two sets. Regard each set as a discrete category. Interpret what a functor $F : X \longrightarrow Y$ means in this case.

(*ii.*) Let $G$ and $H$ be two groups. Regard each group as a one-object category whose morphisms sets correspond to their group elements, with composition their group product. Interpret what a functor $F : G \longrightarrow H$ means in this case.

(*iii.*) Let $X$ and $Y$ be a pair of thin categories. Interpret what a functor $F : X \longrightarrow Y$ means in this case. (Use ($i$) to get you started.)

**2.** Let $G$ be a group. Then for any two elements $a, b \in G$, we define the **commutator** of $a, b$ to be the element

$$aba^{-1}b^{-1}.$$

Define $[G, G]$ to be the set

$$\{x_1 x_2 \cdots x_n \mid n \in \mathbb{N}, x_i \text{ is a commutator in } G\}$$

which we call the **commutator subgroup**. Its underlying set consists of all possible products, with factors that are of the form $a_i b_i a_i^{-1} b_i^{-1}$. One can show that $[G, G] \trianglelefteq G$ for any group $G$, which implies that we may discuss the quotient group $G/[G, G]$, which is abelian in this case.

So, show that we have a functor $F : \mathbf{Grp} \longrightarrow \mathbf{Ab}$ where

$$F(G) = G/[G, G]$$

Deduce the action of $F$ on the morphism of **Grp** (i.e., the group homomorphisms.) and show that it is well-defined.

**3.** Let $R$ be a unital ring. Recall that $GL_n(R)$ is the group consisting of $n \times n$ matrices with entries in $K$. Show that this construction more generally is that of a functor

$$\mathbf{GL}_n : \mathbf{Ring} \longrightarrow \mathbf{Grp}\,.$$

In addition, with such a ring $R$, we may associate it with its group of units $R^\times$, which you may recall is

$$R^\times = \{u \in R \mid ur = ru = 1 \text{ for some } r \in R\}.$$

Show that this also defines a functor

$$(-)^\times : \mathbf{Ring} \longrightarrow \mathbf{Grp}\,.$$

We will see in the next section that there is an interesting relationship between these two functors.

**4.** Recall the category $\mathbf{Set}_{FTO}$ is the category whose objects are sets and whose morphisms are functions with the finite-to-one property (See Exercise 1.3.3). While we saw that $\text{FinC} : \mathbf{Set} \longrightarrow \mathbf{Top}$ where

$$X \mapsto (X, \tau^X_{FC})$$

does **not** define a functor, show that upon changing the domain category from **Set** to $\mathbf{Set}_{FTO}$, it **does** define a functor $\text{FinC} : \mathbf{Set}_{FTO} \longrightarrow \mathbf{Top}$.

**5.** (*i.*) Let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite set. With such a finite set, we can pick a field $k$ and build $X$ into a finite-dimensional vector space $V_X$ over $k$. Explicitly, we can create the vector space

$$V_X = \{c_1 x_1 + \cdots + c_n x_n \mid c_i \in k\}.$$

We define addition in the intuitive way of adding coefficients of the same basis, so this is truly a vector space. Note that when $k = \mathbb{R}$, we get that $V_X \cong \mathbb{R}^n$.

Prove that this process is functorial. That is, show that the functor

$$F : \mathbf{FinSet} \longrightarrow \mathbf{Vect}_k \qquad F(X) = V_X$$

is a functor.

$(ii)$. From any set $X$, we may construct the **free group** $F(X)$ generated by $X$. The elements of $F(X)$ are (1) the elements of $X$, (2) a new element $e$, and (3) all elements $xy$ whenever $x, y \in X$. In this way, $F(X)$ is a group with the product being string concatenation, and we require that $xe = x = ex$. . Below, two words (elements of $F(X)$) are combined.

$$(x^2yz^{-1}) \cdot (zy^2x) = x^2y^2x.$$

Show that we have a functor $F : \mathbf{Set} \longrightarrow \mathbf{Grp}$ where sets are mapped to their free groups, that is, $X \mapsto F(X)$.

$(iii)$. For any set $X$, we can build the **free ring** $(R\{X\}, +, \cdot)$ as follows. Let $(F(X), \cdot)$ be the free group with the added relation that $xy = yx$ for any $x, y \in F(X)$. We can then define

$$R\{X\} = \left\{ \sum_{x_i \in F(X)} x_i^{n_i} \mid \right\}$$

**Note:** This example becomes particularly important later. It can also be generalized to functors $F : \mathbf{Set} \longrightarrow \mathbf{Mon}$, $F : \mathbf{Set} \longrightarrow \mathbf{Ring}$, and other algebraic systems, since sets can also be turned into free monoids, free rings, or other free "objects."

**6.** Let $V$ be a vector space over a field $k$. Recall that we can associate $V$ with its **projective space** $P(V)$ which is defined as the set of equivalence classes of element in $V$, subject to the relation $v \sim w$ if $v = \lambda w$ for some nonzero $\lambda \in k$. That is,

$$P(V) = \Big\{ [v] \mid v \in V \Big\}$$

where $[v]$ denotes the equivalence class of $v$. Show that this process is functorial, so that we have a functor

$$P : \mathbf{Vect}_k \longrightarrow \mathbf{Set}\,.$$

**7.** Let $R$ be a ring with ideal $I$. Recall that we can construct the **radical of the ideal** $I$ as the ideal

$$\sqrt{I} = \{r \in R \mid r^n \in I \text{ for some } n \geq 1\}.$$

Show that we have a functor

$$\sqrt{-} : \mathbf{Ideals}(R) \longrightarrow \mathbf{Ideals}(R)$$

where **Ideals**$(R)$ is the partial order of ideals on $R$, whose ordering is given by subset containment.

**8.** Let $X$ be a topological space, and denote **Open**$(X)$ as the category where the objects are open sets $U \subseteq X$ and morphisms are inclusion morphisms. Create a functor

$$F : \mathbf{Open}(X) \longrightarrow \mathbf{Set}$$

where on objects $F(U) = \{f : U \longrightarrow \mathbb{R} \mid f \text{ is continuous}\}$. That is, how should $F$ act on the morphisms for this to be a functor?

**9.** Let $k$ be a field. With each field, we may associate $k$ with the category $\mathbf{Vect}_k$ which consists of finite dimensional vector spaces $V$ over $k$. Is this process functorial? That is, do we have a functor

$$\mathbf{Vect} : \mathbf{Fld} \longrightarrow \mathbf{Cat}$$

where $\mathbf{Vect}(k) = \mathbf{Vect}_k$?
*Hint: No. But explain why it breaks.*

## 1.8 Forgetful, Full and Faithful Functors.

Like functions, functors can be composed to form new functors.

**Definition 1.8.1.** If $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ are categories where

$$\mathcal{A} \xrightarrow{F} \mathcal{B} \xrightarrow{G} \mathcal{C}$$

are functors, then we can define the **composite functor** $G \circ F : \mathcal{A} \longrightarrow \mathcal{C}$ where

$$C \mapsto G(F(C)) \in \mathcal{C} \qquad (f : A \longrightarrow B) \mapsto G(F(f)) \in \text{Hom}_{\mathcal{C}}(G(F(A)), G(F(B))).$$

We've now reached something quite important. We have the notion of a category, as well as the notion of a functor which acts as a map between categories. Moreover, every category $\mathcal{C}$ is equipped with an identity functor $1_{\mathcal{C}} : \mathcal{C} \longrightarrow \mathcal{C}$, functor composition is associative, and so we may form the **category of categories CAT** where

**Objects.** All categories (large and small)

**Morphisms.** All functors between such categories.

If we instead restrict our objects to all *small* categories, we obtain the category **Cat**, which is usually what we'll work with. Overall, what we see is that functors are the rightful "morphisms" between categories.

Since functors are, in an abstract sense, morphisms, and we know that for general morphisms, there exists a concept of an isomorphism, we can directly apply such a notion to define what an isomorphic functor is.

**Definition 1.8.2.** Let $\mathcal{C}$ and $\mathcal{D}$ be two categories. Then a functor $F : \mathcal{C} \longrightarrow \mathcal{D}$ is said to be a **isomorphism** if it is bijective on both objects and arrows.

Equivalently, $F$ is an isomorphic functor if and only if there exists a functor $G : \mathcal{D} \longrightarrow \mathcal{C}$ such that $F \circ G$ is the identity on $\mathcal{C}$ and $G \circ F$ is the identity on $\mathcal{D}$ (both in terms of objects and arrows).

Sometimes when a functor maps objects from one category to another, the underlying structure of the objects in the first category gets lost. Or perhaps a binary operation acting on the elements in the first set of objects becomes lost. For this, we have a special name.

**Definition 1.8.3.** Let $\mathcal{C}$ and $\mathcal{D}$ be categories and suppose $F : \mathcal{C} \longrightarrow \mathcal{D}$ is a functor. Then $F$ is said to be **forgetful** whenever $F$ does not preserve the axioms and structure present in the objects of $\mathcal{C}$ (whether it be algebraic or some kind of ordering).

The above definition isn't precise, although it is a useful notion to have. It will eventually become precise, but we'll comment more on that after a few examples.

---

**Example 1.8.4.** Consider a group $(G, \cdot)$ with $\cdot$ the binary operation. In some sense, groups are simply sets with added structure, while group homomorphisms are simply functions that respect group structure. Hence we can create a map between **Grp** and **Set** that forgets this structure:

$$(G, \cdot) \mapsto G \qquad \varphi : (G, \cdot) \longrightarrow (H, +) \mapsto \varphi : G \longrightarrow H.$$

We can demonstrate that this process is functorial. Observe that if $1_G : (G, \cdot) \longrightarrow (G, \cdot)$ is the identity group homomorphism, then one can readily note that $1_G(g) = g$ for all $g \in G$, so that it is also an identity function on the underlying set $G$. Therefore, $F(1_G) = 1_{F(G)}$

Next, if $\varphi : G \longrightarrow H$ and $\psi : H \longrightarrow K$ are group homomorphisms, then $F(\psi \circ \varphi)$ is the underlying function $\psi \circ \varphi : G \longrightarrow K$. Note however that for each $g \in G$,

$$F(\psi \circ \varphi)(g) = \psi(\varphi(g)) = F(\psi) \circ F(\varphi)(g) \implies F(\psi \circ \varphi) = F(\psi) \circ F(\varphi).$$

Hence, we see that we have a forgetful functor $F$ : **Grp** $\longrightarrow$ **Set** which leaves behind group operations, and moreover regards every group homomorphism as a function.

---

---

**Example 1.8.5.** Let $(R, +, \cdot)$ be a ring. Recall that $(R, +)$ (alone with its addition) is an abelian group. Hence we can forget the structure of $\cdot : R \times R \longrightarrow R$ and, in a forgetful sense, treat every ring as an abelian group.

This then defines a forgetful functor $F$ : **Rng** $\longrightarrow$ **Ab** which simply maps a ring to its abelian group. This works on the morphisms, since every ring homomorphism $\varphi : (R, +, \cdot) \longrightarrow (S, +, \cdot)$ is a group homomorphism $\varphi : (R, +) \longrightarrow (S, +)$ on the abelian groups.

---

---

**Example 1.8.6.** Consider the category **Top**. Each object in top is a pair $(X, \tau)$ where $\tau$ is a topology on $X$. Moreover, continuous functions are simply functions. This forgetful process is also functorial:

$$(X, \tau) \mapsto X \qquad f : (X, \tau) \longrightarrow (Y, \tau') \mapsto f : X \longrightarrow Y.$$

This then gives us the forgetful functor $F$ : **Top** $\longrightarrow$ **Set**.

---

Some things need to be said about a forgetful functors. You might have noticed that our definition of a forgetful functor was not at all mathematically rigorous. This is because to define forgetful functors we have two main options:

1. Use very deep set theory and logic to characterize the data of a category; then define forgetfulness as forgetting some of the data.
2. Define a forgetful functor to be the left adjoint of a free functor $F : \mathcal{C} \longrightarrow \mathcal{D}$ (usually, $\mathcal{C} =$ **Set**)

Option **1.** sounds like a pain, and I don't know any logic. I'm sure the reader is probably not interested in going on that kind of a ride anyways. Option **2.** is not possible right now, but it will be once we learn about adjunctions.

Thus, using the tools we have right now, we cannot create a *rigorous* mathematical definition of a forgetful functor. This does not mean what we're doing is nonsense; it just means we're being sloppy in the interest of pedagogy. Once we learn about adjunctions things will make more sense, so the reader is urged to not worry too much about the rigor of a forgetful functor.

The sloppiness of our work regarding forgetful functors (i.e., us non-rigorously being like "Hey! See this piece of data? Let's throw it away!") might nevertheless be of some discomfort for the pedantic reader. This is because we cannot rigorously demonstrate what a forgetful functor is at this point; hence a reader interested in true understanding won't be able to fully do so at this point. Sometimes, however, understanding how something works is aided by understanding when something *doesn't* work. Hence to comfort the pedantic reader, we introduce an example where one might intuitively think such a forgetful functor exists, but it in fact does not.

---

**Example 1.8.7.** Recall that the category **hTop** has objects as topological spaces and morphisms as homotopy classes between topological spaces. One might prematurely believe that there is a forgetful functor **hTop** $\longrightarrow$ **Set**, but that is not possible.

In trying to do so, we naturally associate topological spaces $(X, \tau)$ with its underlying set $X$. On morphisms, it's trickier. Suppose $[f : X \longrightarrow Y]$ is a homotopy equivalence class with $f : X \longrightarrow Y$ as the continuous function representing the class. Choose any $f' : X \longrightarrow Y \in [f]$; we may very well choose $f$ itself in which case $f' = f$, and set $F(f') = f'$, where $f' \in \mathbf{Set}$ is regarded as a function.

This breaks when we encounter composition. Suppose $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ are continuous functions. Let $F([f]) = f'$, $G([g]) = g'$, and $F([g \circ f]) = (g \circ f)'$ where $f', g'$, and $(g \circ f)$ are any elements of $[f], [g], [g' \circ f']$ respectively. Then in no case can we always expect that

$$F(g \circ f) = F(g) \circ F(f) \implies (g \circ f)' = g' \circ f'.$$

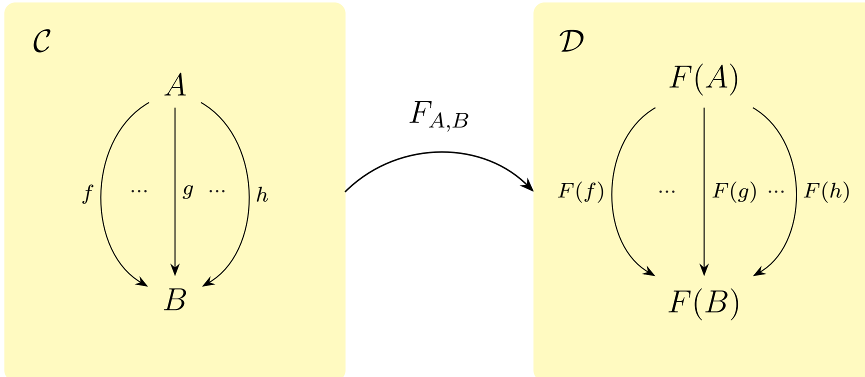
Hence this forgetful process cannot behave functorially.

---

Next, we introduce the notion of **full** and **faithful** functors. Towards that goal, consider a functor $F : \mathcal{C} \longrightarrow \mathcal{D}$ between locally small categories. Then for every pair of objects $A, B \in \mathcal{C}$, there is a function

$$F_{A,B} : \mathrm{Hom}_{\mathcal{C}}(A, B) \longrightarrow \mathrm{Hom}_{\mathcal{D}}(F(A), F(B))$$

where a morphism $f : A \longrightarrow B$ is sent to its image $F(f) : F(A) \longrightarrow F(B)$ under the functor $F$.




As we have a family of functions $F_{A,B}$, we can ask: when is this function surjective or injective? This motivates the following definitions.

**Definition 1.8.8.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor between locally small categories. We say $F$ is

- **Full** if $F_{A,B}$ is surjective
- **Faithful** if $F_{A,B}$ is injective.

If $F_{A,B}$ is an isomorphism, we say $F$ is **fully faithful**.

Now we completely ignored the situation for when $\mathcal{C}, \mathcal{D}$ are not locally small. This is out of pedagogical interest; if $\mathcal{C}, \mathcal{D}$ are not locally small then we do not have the function described above. However, the concept of full and faithful can still be described; it's just not as nice of a description as before.

**Definition 1.8.9.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor.

- **Full** if for all $A, B$, every morphism $g : F(A) \longrightarrow F(B)$ in $\mathcal{D}$ is the image of some $f : A \longrightarrow B$ in $\mathcal{C}$
- **Faithful** if for all $A, B$, we have that if $f_1, f_2 : A \longrightarrow B$ with $F(f_1) = F(f_2)$, then$f_1 = f_2$.

We then say $F$ is a **fully faithful** if it is both full and faithful.

---

**Example 1.8.10.** Consider the forgetful functor $F : \textbf{Top} \longrightarrow \textbf{Set}$ which we introduced earlier; topological spaces $(X, \tau)$ are sent to their underlying sets $X$ while continuous functions $f : (X, \tau) \longrightarrow (Y, \tau')$ are regarded as functions $f : X \longrightarrow Y$. This functor is faithful, since if two continuous functions are equal as set maps, then they are equal as continuous functions. The fact that this functor is faithful is simply due to the fact that the extra data on a continuous function, i.e., its continuity, does not interfere with its behavior of being a set function in sending points $X$ to $Y$.

Note however that this function is clearly not full, because not every function $g : X \longrightarrow Y$ can be regarded as a continuous function between the topological spaces.

---

---

**Example 1.8.11.** Let $(G, \cdot)$ and $(H, \cdot)$ be a group. Regard both groups as one object categories

$\mathcal{C}$ and $\mathcal{D}$ with objects $\bullet$ and $\bullet$ where we set

$$\text{Hom}_{\mathcal{C}}(\bullet, \bullet) = G \qquad \text{Hom}_{\mathcal{C}}(\bullet, \bullet) = H$$

so that each $g \in G$ is now a morphism $g : \bullet \longrightarrow \bullet$, and vice versa for every $h \in G$, so that composition is given by the group structure. If we have a functor $F : \mathcal{C} \longrightarrow \mathcal{D}$ between these categories, then the function we introduced simply becomes a set function

$$F_{\bullet,\bullet} : \text{Hom}_{\mathcal{C}}(\bullet, \bullet) \longrightarrow \text{Hom}_{\mathcal{D}}(\bullet, \bullet).$$

However, the functorial properties allow this to extend to a group homomorphism from $G$ to $H$. Therefore, we see that if $F : \mathcal{C} \longrightarrow \mathcal{D}$ is full, it extends to a surjective group homomorphism. If it is faithful, it extends to an injective group homomorphism.

---

---

**Example 1.8.12.** Consider the category of **Grp**, and recall it has a forgetful functor $F$ : **Grp** $\longrightarrow$ **Set**. This functor is actually fully faithful; to see this, consider two group homomorphisms $\varphi, \psi : (G, \cdot) \longrightarrow (H, \cdot)$, and suppose that $F(\varphi) = F(\psi)$. Then this implies that $F(\varphi)(g) = F(\psi)(g)$ for each $g \in G$. However, $F(\varphi)(g) = \varphi(g)$ and vice versa for $\psi$. Therefore, we have that $\varphi = \psi$, so that the forgetful functor $F$ is a faithful functor.

---

The above example can be repeated for many familiar categories, which motivates the following definition.

**Definition 1.8.13.** A category $\mathcal{C}$ is said to be **concrete** if there is a faithful functor $F$ : $\mathcal{C} \longrightarrow$ **Set**.

Examples of concrete categories includ **Grp**, **Top**, $R$ mod , and many others since these categories are, in some sense, built from **Set**. Their objects are sets, and their morphisms are functions with extra properties; nevertheless, at the end of the day the morphisms are still functions. Note in particular that these categories are not subcategories of **Set**, but they are still deeply related to this category in a way that the above definition illuminates.

We don't have the tools right now, but we will later show that every small category $\mathcal{C}$ is a concrete category.

## Exercises

1. In this exercise, you'll demonstrate that the image of a functor is generally not a category, but that full functors remedy the situation.

   (*i.*) Let $F : \mathcal{C} \longrightarrow \mathcal{D}$. Define the **image of** $F$ in $\mathcal{D}$ to consist of

   **Objects.** All $F(A)$ with $A \in \mathcal{C}$

**Morphisms.** For any two objects $F(A)$ and $F(B)$, we have that

$$\mathrm{Hom}_{\mathcal{D}}(F(A), F(B)) = \{F(f) \mid f : A \longrightarrow B\}.$$

Show that this is not always a category. In general, the image of a functor is not a category.
*Hint*: Picture two categories $\mathcal{C}$ and $\mathcal{D}$ below

$$\mathcal{C}: \quad \begin{tikzcd} A \arrow[d, "f_1"'] & C \arrow[d, "f_2"'] \\ B & D \end{tikzcd} \qquad\qquad \mathcal{D}: \quad \begin{tikzcd} X \arrow[rr, "g_3"] \arrow[dr, "g_1"'] & & Z \\ & Y \arrow[ur, "g_2"'] & \end{tikzcd}$$

and consider the functor $F(A) = X, F(B) = F(C) = Y$, and $F(D) = Z$. Explain what goes wrong, and more generally why the image of a functor is not a category.

(*ii.*) Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a full functor. Show that the image of $\mathcal{C}$ under $F$ forms a full subcategory of $\mathcal{D}$.

(*iii.*) By (*ii*), it is sufficient for $F$ to be full in order for the image to be a category. Is this condition *necessary* for the image to form a category? In other words, suppose the image of a functor $F$ is a category. Is $F$ full?

# 1.9 Natural Transformations

**Example 1.9.1.** Suppose we have a pair of functors $F, G : \mathcal{C} \longrightarrow \textbf{Set}$. In particular, suppose that $F(A) \subseteq G(A)$ for all objects $A$. This means that for each $A$, there exists an injection $i_A : F(A) \longrightarrow G(A)$.

Now this is a bit of an interesting construction since for any morphism $f : A \longrightarrow B$ in $\mathcal{C}$, there are now two ways we can get from $F(A)$ to $G(B)$.

$$
\begin{tikzcd}
A \arrow[d, "f"] \\
B
\end{tikzcd}
\qquad
\begin{tikzcd}
F(A) \arrow[r, "i_A"] \arrow[d, "F(f)"'] & G(A) \arrow[d, "G(f)"] \\
F(B) \arrow[r, "i_B"'] & G(B)
\end{tikzcd}
\qquad
\begin{tikzcd}
x \arrow[r, mapsto] \arrow[d, mapsto] & x \arrow[d, mapsto] \\
F(f)(x) \arrow[r, mapsto] & F(f)(x) \stackrel{?}{=} G(f)(x)
\end{tikzcd}
$$

As we have two different ways of traversing this diagram, **are they equivalent**? That is, is it the case that

$$G(f) \circ i_A = i_B \circ F(f) \quad \text{or, spelled out,} \quad F(f)(x) = G(f)(x)?$$

In general, this isn't true. But one way (and as we'll see in the future, the *only* way) we can make this diagram commute is if

$$F(f) = G(f)\Big|_{F(A)}.$$

That is, if $F(f)$ is a restriction of $G(f)$. We summarize this observation by stating that, if $F(f) = G(f)\Big|_{F(A)}$ for all $f$, then the inclusion $i_A : F(A) \longrightarrow G(A)$ is *natural*.

---

**Example 1.9.2.** Let $X$ be a topological space. Then we can create the abelian groups

$$C_0(X), C_1(X), \dots, C_n(X), \dots$$

Here, $C_n(X)$ is the free abelian group generated by continuous functions of the form $\varphi : \Delta^n \longrightarrow X$, where where $\Delta^n$ is the $n$-simplex. Hence, elements are of $C_n$ are of the form

$$\sum_{\varphi} n_{\varphi} \cdot \varphi$$

where all but finitely many of the integer coefficients $n_\varphi$ are zero.

In algebraic topology, one observes that these abelian groups assemble into a chain via a boundary operator $\partial_n : C_n \longrightarrow C_{n-1}$ with the property that $\partial_{n+1} \circ \partial_n = 0$ for all $n$.

$$\cdots \xrightarrow{\partial_{n+1}} C_n(X) \xrightarrow{\partial_n} C_{n-1}(X) \xrightarrow{\partial_{n-1}} \cdots \xrightarrow{\partial_1} C_0(X)$$

Now suppose that $f : X \longrightarrow Y$ is a continuous map between topological spaces. Then for each $n$, there is an evident mapping between the chain complexes.

$$C_n(f) : C_n(X) \longrightarrow C_n(Y) \qquad \sum_{\varphi} n_\varphi \cdot \varphi \mapsto \sum_{\varphi} n_\varphi \cdot f \circ \varphi.$$

This is because if $\varphi : \Delta^n \longrightarrow X$ is a singular map then $f \circ \varphi : \Delta^n \longrightarrow Y$ is also a singular map because $f$ is continuous. However this presents us with an issue, one we faced in the earlier example. On one hand, we have a map $C_{n-1}(f) \circ \partial_n : C_n(X) \longrightarrow C_n(Y)$. On the other hand, we have a map $\partial_n \circ C_n(f) : C_n(X) \longrightarrow C_n(Y)$. But are these equivalent maps?

$$\begin{tikzcd}
C_n(X) \arrow[r, "\partial_n"] \arrow[d, "C_n(f)"'] & C_{n-1}(X) \arrow[d, "C_{n-1}(f)"] \\
C_n(Y) \arrow[r, "\partial_n"'] & C_{n-1}(Y)
\end{tikzcd}$$

It's a simple exercise to show that this diagram does in fact commute, i.e., that $C_{n-1}(f) \circ \partial_n = \partial_n \circ C_n(f)$ for all $n$.

As a result, this "natural" result (again pun intended) gives us intuition on how to define a mapping between two chain complexes $\{C_n\}_{n\in\mathbb{N}}$ and $\{C_n\}_{n\in\mathbb{N}}$: : it is any family of maps $\psi_n : C_n \longrightarrow C'_n$ such that $\psi_{n-1} \circ \partial_n = \partial_n \circ \psi_n$. Moreover, since we have a notion of objects (i.e, chain complexes $\{C_n\}$ ) and morphisms (chain maps) this gives rise to a category **Ch(Ab)**, the category of chain complexes of abelian groups.

---

When the two ways to traverse the diagram are equivalent, we call this behavior **natural** and it makes mathematicians very happy. Naturality, which is what we will refer to this property as, is ubiquitous in mathematics and functors give us a convenient way of conceptualizing this useful property.

**Definition 1.9.3.** Let $F, G : \mathcal{C} \longrightarrow \mathcal{D}$ be two functors. Then we define a mapping[7] between the functors

$$\eta : F \longrightarrow G$$

to be a **natural transformation** if it associates each $C \in \text{Ob}(\mathcal{C})$ with a morphism

$$\eta_C : F(C) \longrightarrow G(C)$$

in $\mathcal{D}$ such that for every $f : A \longrightarrow B$, we have that

[7]Think **morphism**, because the word mapping here doesn't rigorously mean anything. That's because we don't really have a word to describe what a natural transformation really is. We have axioms, which we present, but we don't have a nice word. That nice word will turn out to be morphism, and you will see soon why.

$$
\begin{tikzcd}
A \arrow[d, "f"] \\
B
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
F(A) \arrow[r, "\eta_A"] \arrow[d, "F(f)"'] & G(A) \arrow[d, "G(f)"] \\
F(B) \arrow[r, "\eta_B"'] & G(B)
\end{tikzcd}
$$

which amounts to $\eta_B \circ F(f) = G(f) \circ \eta_A$.

Thus we can imagine that $\eta$ translates the diagram produced by the functor $F$ to a diagram produced by $G$. For example; if $\eta$ is a natural transformation between $F$ and $G$, then we also see that the following diagram commutes:

$$
\begin{tikzcd}
A \arrow[dr, "f"] \arrow[dd, "h"'] & \\
& B \arrow[dl, "g"] \\
C &
\end{tikzcd}
\qquad
\begin{tikzcd}
F(A) \arrow[rrr, "\eta_A"] \arrow[dr, "F(f)"] \arrow[dd, "F(h)"'] & & & G(A) \arrow[dr, "G(f)"] \arrow[dd, "G(h)"] & \\
& F(B) \arrow[rrr, "\eta_B"] \arrow[dl, "F(g)"] & & & G(B) \arrow[dl, "G(g)"] \\
F(C) \arrow[rrr, "\eta_C"] & & & G(C) &
\end{tikzcd}
$$

and this diagram commutes

$$
\begin{tikzcd}
A \arrow[r, "f", green] \arrow[d, "h"', red] & B \arrow[d, "g", blue] \\
C \arrow[r, "k"', orange] & D
\end{tikzcd}
\qquad
\begin{tikzcd}
& F(A) \arrow[rr] \arrow[dl, green] \arrow[dd, red] & & G(A) \arrow[dl, green] \arrow[dd, red] \\
F(B) \arrow[rr] \arrow[dd, blue] & & G(B) \arrow[dd, blue] & \\
& F(C) \arrow[rr] \arrow[dl, orange] & & G(C) \arrow[dl, orange] \\
F(D) \arrow[rr] & & G(D) &
\end{tikzcd}
$$

if the above diagram on the left commutes. Colors are added to aid the visualization in seeing how the natural transformation translates the diagram produced by $F$ to the diagram produced by $G$.

**Definition 1.9.4.** Let $\eta : F \longrightarrow G$ be a natural transformation. If $\eta_A : F(A) \longrightarrow G(A)$ is an isomorphism for each object $A$, then we say $\eta$ is a **natural isomorphism**.

---

**Example 1.9.5.** Let $K$ be a ring in **CRng**. Recall from Exercise 1.3.3 that

$$GL_n(-) : \mathbf{CRing} \longrightarrow \mathbf{Grp} \quad (-)^{\times} : \mathbf{CRing} \longrightarrow \mathbf{Grp}$$

are functors. In that exercise we actually showed that the domain categories were **Ring**, but for our purpose we can restrict these functors to the full subcategory **CRing**.

Consider a commutative ring $K$. Recall that for matrix $M \in GL_n(K)$, we can take the determinant of $K$; we are usually more familiar with this concept when $K = \mathbb{R}$. However, it is a fact from ring theory that a matrix $M$ is invertible if and only if the determinant $\det(M)$ of $M$ is in $K^\times$. Since $GL_n(K)$ is the set of all such invertible matrices, we see that we may associate each $K$ with its determinant function

$$\det_K : GL_n(K) \longrightarrow K^\times$$

which sends an invertible $M \in GL_n(K)$ to its determinant in $K^\times$. To see that this morphism is a group homomorphism, we simply recall the determinant property

$$\det(AB) = \det(A)\det(B).$$

The claim is now that this family of morphisms assembles into a natural transformation. Specifically, that $\det : GL_n(-) \longrightarrow (-)^\times$. To see, this, let $f : K \longrightarrow K'$ be a homomorphism between commutative rings. Recall from ring theory that the determinant of a matrix $M = [a_{ij}]$ with $a_{ij} \in K$ is given by

$$\det(M) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

where $S_n$ is the symmetric group, and $\operatorname{sgn}(\sigma)$ is the sign of a permutation. Now for det to form a natural transformation, we'll need that the diagram below commutes.

$$\begin{tikzcd} K \arrow[d, "f"] \\ K' \end{tikzcd} \qquad \begin{tikzcd} GL_n(K) \arrow[r, "\det_K"] \arrow[d, "GL_n(f)"'] & K^\times \arrow[d, "f^\times"] \\ GL_n(K') \arrow[r, "\det_{K'}"'] & K'^\times \end{tikzcd}$$

Note that $f : K \longrightarrow K'$ is a commutative ring homomorphism. To show this diagram commutes, consider any $M = [a_{ij}] \in GL_n(K)$. Observe that

$$\begin{aligned} (f^\times \circ \det_K)(M) &= f^\times(\det_K(M)) \\ &= f^\times\left(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}\right) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) f(a_{1\sigma(1)}) \cdots f(a_{n\sigma(n)}) \\ &= \det_{K'}([f(a_{ij}]]) \\ &= \det_{K'} \circ GL_n(f)(M). \end{aligned}$$

Hence we see that the diagram commutes, so that the determinant $\det : GL_n(-) \longrightarrow (-)^\times$ assembles into a natural transformation between the functors.

---

---

**Example 1.9.6.** For a field $k$, recall that we have two functors $A^n(-), P^n(-) : \mathbf{Fld} \longrightarrow \mathbf{Set}$ where

$$A^n(k) = \{(a_0, \dots, a_{n-1}) \mid a_i \in k\} \qquad P^n(k) = A^{n+1}(k)/\sim$$

where $\sim$ is the equivalence relation on the set $A^{n+1}(k)$ described as follows: $(a_0, \dots, a_n) \sim (a'_0, \dots, a'_n)$ if $(a_0, \dots, a_n) = \lambda(a'_0, \dots, a'_n)$ for some nonzero $\lambda \in k$. Geometrically, the equivalence relation identifies points which are lying on the same line passing through the origin.

As we noted before, these functors are particularly important in algebraic geometry. Now for each point $(a_0, \dots, a_n)$, denote $[(a_0, \dots, a_n)]$ as its equivalence class. Let $\theta_k : A^{n+1}(k) \longrightarrow P^n(k)$ be the function that maps a point $(a_0, \dots, a_n)$ to its equivalence class $[(a_0, \dots, a_n)]$. Our claim is that for each $k$, the functions $\theta_k$ assemble into a natural transformation.

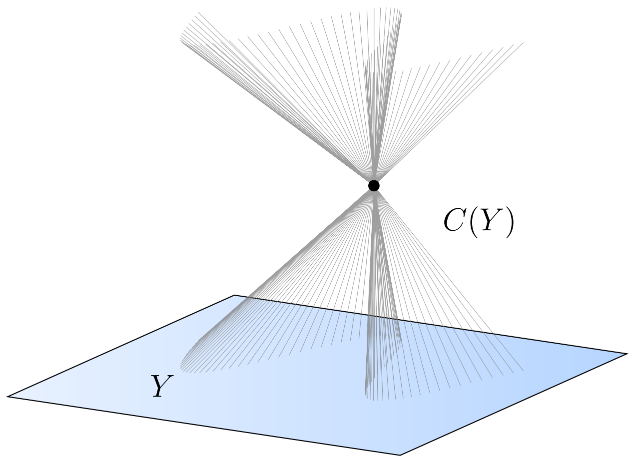
That is, for a field homomorphism $\varphi : k \longrightarrow k'$, the diagram

$$\begin{tikzcd} k \arrow[d, "\varphi"] & A^{n+1}(k) \arrow[r, "\theta_k"] \arrow[d, "A^{n+1}(\varphi)"'] & P^n(k) \arrow[d, "P^n(\varphi)"] \\ k' & A^{n+1}(k') \arrow[r, "\theta_{k'}"'] & P^n(k') \end{tikzcd}$$

commutes. The reader is encouraged to fill in the details for this one. It's quite surprising that this does assemble into a natural transformation, because in general there is no reason to ever expect that the projection map, $\pi : X \longrightarrow X/\sim$ with $\sim$ an equivalence relation, is, in any sense, natural. Its because most functions mess things up, and disorganize the equivalence classes!

The above morphism, $\theta : A^{n+1} \longrightarrow P^n$, actually has a very interesting geometric realization[8]. If $Y$ is an algebraic subset of $P^n(k)$, then we can build the **affine cone** $C(Y) = \theta^{-1}(Y) \cup \{(0, \dots, 0)\}$. With $n = 2$, $Y$ corresponds to a curve in $P^2(k)$, which generates the surface $C(Y)$ in in $A^3(k)$.

[8]This isn't important for the reader to understand. However, I do want to avoid blabbering abstract nonsense so that the reader knows we're doing real, relevant mathematics. And perhaps it might be motivation for the reader to check out an algebraic geometry text!




---

---

**Example 1.9.7.** Earlier, we showed that $p_G : \mathbf{Grp} \longrightarrow \mathbf{Ab}$ in which $G \mapsto G/[G,G]$ was a functor. It turns out that the projection

$$T_G : G \longrightarrow G/[G,G] \qquad g \mapsto g + [G,G]$$

forms a natural transformation between the identity functor $1_{\mathbf{Grp}} : \mathbf{Grp} \longrightarrow \mathbf{Grp}$ on $\mathbf{Grp}$ and the functor $p_G$.

To show this, consider the morphism $f : G \longrightarrow H$ in $\mathbf{Grp}$. We know that $p_G$ induces a function $f^* : G/[G,G] \longrightarrow H/[H,H]$ defined as

$$f^*(g + [G,G]) = f(g) + [H,H].$$

Now let $g \in G$.

$\boldsymbol{T_H \circ f(g)}$**.** On one hand, observe that

$$T_H \circ (f(g)) = f(g) + [H,H].$$

$\boldsymbol{f^* \circ (T_G(g))}$**.** On the other hand, observe that

$$f^* \circ T_G(g) = f^*(g + [G,G]) = f(g) + [H,H].$$

Hence, we see that

$$T_H \circ f = f^* \circ T_G$$

so that the following diagram commutes

$$\begin{tikzcd} G \arrow[r, "T_G"] \arrow[d, "f"'] & G/[G,G] \arrow[d, "f^*"] \\ H \arrow[r, "T_H"] & H/[H,H] \end{tikzcd}$$

and hence $T$ is a natural transformation.

---

**Example 1.9.8.** The categories **FinOrd** and $Set_{\boldsymbol{F}}$, are closely related categories. Recall that **FinOrd** has finite ordinals $n = \{0, 1, 2, \dots, n-1\}$ as objects with morphisms all functions $f : m \longrightarrow n$ where $m, n$ are natural numbers, and the objects of $\mathbf{Set}_{\boldsymbol{F}}$ are all finite sets (of some universe $U$) with morphisms all functions between such sets.

Obviously the objects and morphisms of **FinOrd** are in $\mathbf{Set}_{\boldsymbol{F}}$. Thus, let $S : \mathbf{Findord} \longrightarrow \mathbf{Set}_{\boldsymbol{F}}$ be the inclusion functor.

Define a functor $\# : \mathbf{Set}_{\boldsymbol{F}} \longrightarrow \mathbf{FinOrd}$ as follows. Assign each $X \in \mathbf{Set}_{\boldsymbol{F}}$ to the ordinal $\#X = n$, the number of elements in $X$. We can represent this bijective mapping as

$$\theta_X : X \longrightarrow \#X.$$

Furthermore, if $f : X \longrightarrow Y$ is a morphism in $\mathbf{Set}_{\boldsymbol{F}}$, associate $f$ with the morphism $\#f : \#X \longrightarrow \#Y$ in **FinOrd** defined by

$$\#f = \theta_Y \circ f \circ \theta_X^{-1}.$$

Thus we have that the following diagram is commutative:

$$\begin{tikzcd} X \arrow[r, "\theta_X"] \arrow[d, "f"'] & \#X \arrow[d, "\#f"] \\ Y \arrow[r, "\theta_Y"] & \#Y \end{tikzcd}$$

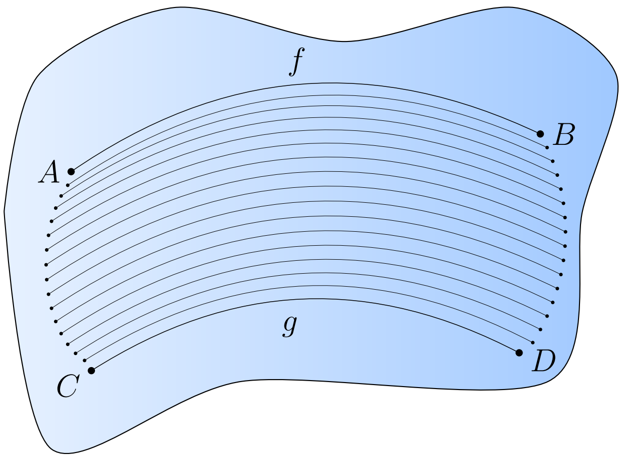
and $\theta$ acts a natural transformation between the two functors.

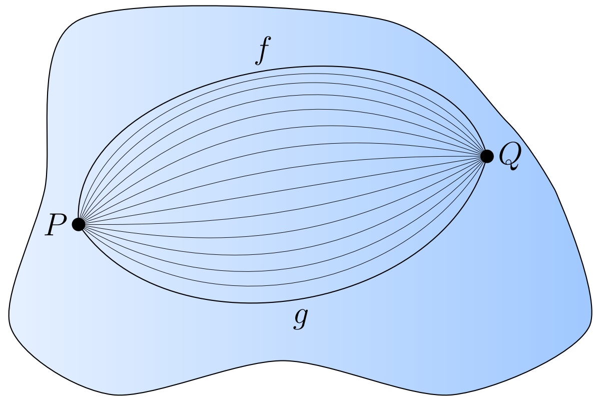
Note that if $X$ is an ordinal number, we define $\theta_X$ to be the identity function, which ensures that $\# \circ S$ is the identity functor on **FinOrd**. However, $S \circ \#$ is not the identity on $\mathbf{Set}_{\boldsymbol{F}}$, since the input will be $X$ while the output will just be $\#X$ (as $S$ is just the inclusion functor.)

---

To end this section, we offer a topological interpretation of the concept of a natural transformation, one which has been known by category theorists since the 1960's, but a perspective which usually is not introduced since it does not really offer signficant pedagogical advantagous unless the reader is already aware of basic homotopy theory (in which case, they probably already know what a natural transformation is). I've nevertheless decided to include it because it is an interesting perspective.

Let $X$ and $Y$ be topological spaces. Consider two functions $f : X \longrightarrow Y$. Recall that a **homotopy** $H$ from $X$ to $Y$ is a continuous function $H : [0,1] \times X \longrightarrow Y$ such that $H(0,x) = f(x)$ and $H(1,x) = g(x)$. A simple example of a homotopy is when $X = [0,1]$. In this case, $f, g : [0,1] \longrightarrow Y$ are simply two continuous paths in $Y$. A homotopy, in this situation, between $f, g$ is pictured on the bottom left.

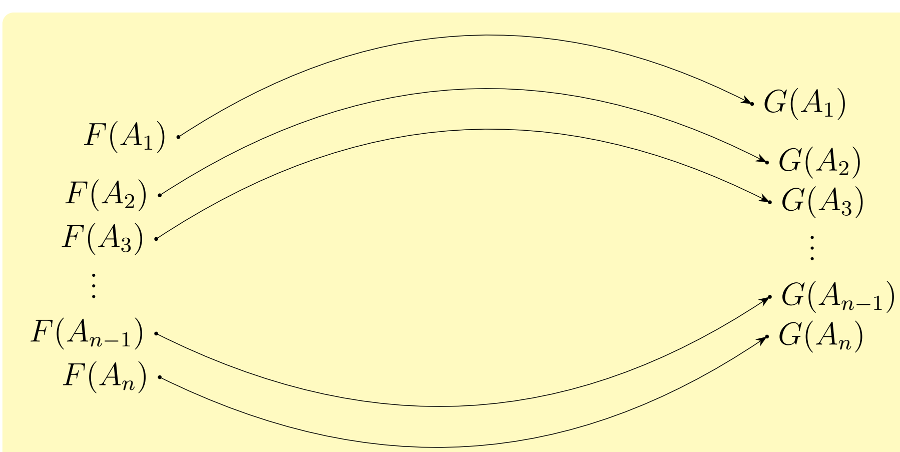



On the above right we have the situation for when $f, g$ start and end at the same point; this homotopy is know as a **path homotopy**.

Of course, a homotopy doesn't always exist. When it does, a homotopy can be interpreted as parameterizing, via $t \in [0,1]$, a family of continuous functions $H_t : X \longrightarrow Y$ which *continuously* deform $f$ into $g$[9].

But this story is familar! A natural transformation $\eta : F \longrightarrow G$ between two functors $F, G : \mathcal{C} \longrightarrow \mathcal{D}$ give rise to a family of morphisms $\eta_A : F(A) \longrightarrow G(A)$ which are parameterized by the objects of $\mathcal{C}$ (which also satisfy the naturality property). Below we have this pictured of what this generally looks like.




[9]Caution: a family of continuous functions does not conversely define a homotopy.

So, what gives? Is the concept of a natural transformation somewhat logically and conceptually analogous to the concept of a homotopy? The answer is yes, and we can define a natural transformation in the following manner which is strikingly similar to the definition of a homotopy.

**Definition 1.9.9.** Let $F, G : \mathcal{C} \longrightarrow \mathcal{D}$ be functors. Let $\mathbf{2}$ be the category with two objects $0, 1$ and a single nontrivial morphism. A **natural transformation** $\eta : F \longrightarrow G$ is a functor $\eta : \mathcal{C} \times (\mathbf{2}) \longrightarrow \mathcal{D}$ such that $\eta(-, 0) = F$ and $\eta(-, 1) = G$.

Proving this is left as an exercise.

## Exercises

**1.** In what follows, let $F, G : \mathcal{C} \longrightarrow \mathcal{D}$ be a pair of functors. Interpret what a natural transformation $\eta : F \longrightarrow G$ is in each case.

(*i.*) Where $\mathcal{C}$ is a discrete category, and $\mathcal{D}$ is arbitrary. Separately, can we have a natural transformation when $\mathcal{D}$ is discrete?

(*ii.*) Where $\mathcal{C}$ and $\mathcal{D}$ are preorders.

(*iii.*) Where $\mathcal{C}$ and $\mathcal{D}$ are one-object categories whose morphisms are group.

(*iv.*) Where $\mathcal{C}$ is arbitrary and $\mathcal{D}$ is **Cat**.

**2.** Show that Definition 1.9.9 and Definition 1.9.3 are equivalent.

**3.** Consider the initial discussion of this section. Prove that for two functors $F, G : \mathcal{C} \longrightarrow \mathbf{Set}$ such that $F(A) \subseteq G(A)$ for all $A \in \mathcal{C}$, the inclusion morphisms $i_A : F(A) \longrightarrow G(A)$ form a natural transformation $i : F \longrightarrow G$ if and only if, for each $f : A \longrightarrow B$ in $\mathcal{C}$, we have that $F(f) = G(f)|_{F(A)}$.

**4.** Let $\mathcal{C}$ be a category, and consider two objects $A, B$ so that we have the functors

$$\mathrm{Hom}_{\mathcal{C}}(A, -), \mathrm{Hom}_{\mathcal{C}}(B, -) : \mathcal{C} \longrightarrow \mathbf{Set}\,.$$

(*i.*) Let $\varphi \in \mathrm{Hom}_{\mathcal{C}}(B, A)$. Show that the family of functions

$$\varphi^*_C : \mathrm{Hom}_{\mathcal{C}}(A, C) \longrightarrow \mathrm{Hom}_{\mathcal{C}}(B, C)$$

indexed by each object $C \in \mathcal{C}$, where $\varphi^*_C(f : A \longrightarrow C) = f \circ \varphi : B \longrightarrow C$, forms a natural transformation $\varphi^* : \mathrm{Hom}_{\mathcal{C}}(A, -) \longrightarrow \mathrm{Hom}_{\mathcal{C}}(B, -)$.

(*ii.*) Show that every natural transformation $\eta : \mathrm{Hom}_{\mathcal{C}}(A, -) \longrightarrow \mathrm{Hom}_{\mathcal{C}}(B, -)$ is constructed in this way.

**5.** Let $F : \mathcal{C} \longrightarrow \mathbf{Set}$ be any other functor. Interpret what a natural transformation $\eta : \bullet \longrightarrow F$ is. What about $\varepsilon : F \longrightarrow \bullet$?

**6.** For every ring $R$ there is a natural inclusion homomorphism $i_R : R \longrightarrow R[x]$. Thus, let

$(-)[x] : \textbf{Ring} \longrightarrow \textbf{Ring}$ be the functor that sends a ring $R$ to its single-variable polynomial ring $R[x]$. Show that we have a natural transformation

$$i : I \longrightarrow (-)[X]$$

where $I : \textbf{Ring} \longrightarrow \textbf{Ring}$ is the identity on **Ring**.

**7.** Recall the category of $G$-sets is the category where

**Objects.** All $G$-sets $X$ (i.e., sets $X$ such that $G$ has a group action $\varphi : X \times G \longrightarrow X$)
**Morphisms.** All $G$-equivariant morphisms (i.e., functions $f : X \longrightarrow Y$ such that $f(g \cdot x) = g \cdot f(x)$).

(Also see Exercise 1.3.6). Let $X$ be a $G$-set with action map $\varphi : X \times G \longrightarrow X$ and fix an element $g \in G$. For such an $X$, define the map $\varphi_X^g : X \longrightarrow X$ where $\varphi_X^g(x) = \varphi(g, x)$.

Show that for each $g$, the maps $\varphi^g$ form a natural transformation $I \longrightarrow I$, where $I : \textbf{G-sets} \longrightarrow \textbf{G-sets}$ is the identity functor on this category. (Note that this is a nontrivial example of a natural transformation between a functor and itself!)

# 1.10 Monic, Epics, and Isomorphisms

In category theory the ultimate focus is placed on the morphisms within a category. What we really care about are the relationships between the objects. Thus in this section we'll go over *types* of morphisms that exist between objects.

The way that this is done in set theory is to consider injective functions, surjective functions, and isomorphisms. This can also be done in topology, and in group, ring, and module theory. However, these concepts make no sense in general. This is because in general, the morphisms of a category are not functions because in general, the objects of a category are not sets (even if the objects are sets, the morphisms can still be different than functions).

We can nevertheless abstract the concept of injections and surjections by expressing their properties categorically; that is, without reference to specific elements in any objects. This leads to the concepts of monomorphisms and epimorphisms.

**Definition 1.10.1.** Let $f : A \longrightarrow B$ be a morphism. Then

1. $f$ is a **monomorphism** (or is monic) if

$$f \circ g_1 = f \circ g_2 \implies g_1 = g_2$$

for all $g_1, g_2 : C \longrightarrow A$, where $D$ is arbitrary.

$$\begin{tikzcd} C \arrow[r, shift left, "g_1"] \arrow[r, shift right, "g_2"'] \arrow[dr, "f\circ g_1 = f\circ g_2"'] & A \arrow[d, "f"] \\ & B \end{tikzcd}$$

2. $f$ is a **epimorphism** (or is epic) if

$$g_1 \circ f = g_2 \circ f \implies g_1 = g_2$$

for all $g_1, g_2 : B \longrightarrow C$, where $C$ is an arbitrary object.

$$\begin{tikzcd} A \arrow[d, "f"'] \arrow[dr, "g_1\circ f = g_2\circ f"] & \\ B \arrow[r, shift left, "g_1"] \arrow[r, shift right, "g_2"'] & C \end{tikzcd}$$

3. $f$ is a **split monomorphism** (or retraction) if, for some $g : B \longrightarrow A$,

$$f \circ g = 1_B.$$

4. $f$ is a **split epimorphism** (or section) if, for some $g : B \longrightarrow A$,

$$g \circ f = 1_A.$$

Monomorphisms and epimorphisms are an abstraction that take advantage key properties of both injective and surjective functions. We illustrate this with a few examples.

---

**Example 1.10.2.** In **Set**, an injective function $f : X \longrightarrow Y$ is "one-to-one" in the sense that $f(x) = f(y)$ if and only if $x = y$. With that said, suppose that $g_1, g_2 : Z \longrightarrow X$ are functions and moreover that $f \circ g_1 = f \circ g_2$. Then this means that, for all $z \in Z$, we have that

$$f(g_1(z)) = f(g_2(z)) \implies g_1(z) = g_2(z)$$

since $f$ is one-to-one. Hence we see that injective functions are monomorphisms in **Set**; one can then conversely show that a monomorphism in **Set** are injective functions.

---

**Example 1.10.3.** Let $(G, \cdot)$ be a group, and suppose $(H, \cdot)$ is a normal subgroup of $G$. Then with such a construction, we always have access to the inclusion and projection homomorphisms

$$i : H \longrightarrow G \qquad i(h) = h$$
$$\pi : G \longrightarrow G/H \qquad \pi(g) = g + H.$$

It is not hard to see that $i$ is a monomorphism and $\pi$ is an epimorphism; for suppose $\varphi, \psi : K \longrightarrow G$ are two group homomorphisms from some group $K$ where $i \circ \varphi = i \circ \psi$. Then for each $k \in K$, $i(\varphi(k)) = i(\psi(k)) \implies \varphi(k) = \psi(k)$, so that $\varphi = \psi$. Conversely, if $\sigma, \tau : G \longrightarrow M$ are two group homomorphisms to some group $M$ such that $\sigma \circ \pi = \tau \circ \pi$, then because $\pi$ is surjective we have that $\sigma = \tau$. Hence, we see $\pi$ is an epimorphism.

Since the above constructions can be repeated in the categories **Ab**, **Ring**, and $R$ mod , so can the above argument. We'll see more generally the deeper reason for why this is the case later on.

---

**Example 1.10.4.** In the category of fields, **Fld**, every nonzero morphism is a monomorphism. This is due to the classic argument: the only nontrivial ideal of a field $k$ its itself; hence the kernal of any map $\varphi : k \longrightarrow k'$ is either trivial or all of $k$. If we suppose $\varphi$ is nonzero, then we see that it must be injective, and hence a monomorphism.

---

**Definition 1.10.5.** Let $f : A \longrightarrow B$ be a morphism between two objects $A$ and $B$. We say that $f$ is an **isomorphism** if there exists a morphism $f^{-1} : B \longrightarrow A$ in $\mathcal{C}$! such that

$$f \circ f^{-1} = \text{id}_A \qquad f^{-1} \circ f = \text{id}_B .$$

In this case, $f^{-1}$ is unique, and for any two isomorphisms $f : A \longrightarrow B$ and $g : B \longrightarrow C$ we have

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

In this case we say that $A$ and $B$ are isomorphic and denote this as $A \cong B$.

This is a generalization of the familiar concept of isomorphisms in abstract algebra and in set theory that one usually encounters.

Next, we illustrate a few properties of these types of morphisms.

**Proposition 1.10.6.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor. Then if $f : A \longrightarrow B \in \mathcal{C}$

- is an isomorphism, then $F(f)$ is an isomorphism in $\mathcal{D}$.
- is a split monomorphism, then $F(f)$ is a split monomorphism in $F(f)$
- is a split epimorphism, then $F(f)$ is a split epimorphism.

That is, functors **preserve** isomorphisms, split monomorphism, and split epimorphisms.

In general, functors do not **reflect** isomorphisms, split monomorphisms, and split epimorphisms. That is, if $F(f) : F(A) \longrightarrow F(B)$ is an isomorphism it is not the case that $f$ is an isomorphism.

We demonstrate this with the following example.

---

**Example 1.10.7.** Recall that $\text{Spec}(-) : \textbf{CRing} \longrightarrow \textbf{Set}$ is a functor that appears in algebraic geometry. It sends every commutative ring $A$ to its ring spectrum $\text{Spec}(A)$, which consists of all prime ideals of $A$.

Let $N = \bigcap_{P \in \text{Spec}(A)} P$ be the intersection of all prime ideals. An equivalent way to speak of $N$ is the set $N = \{a \in A \mid a^m = 0 \text{ for some positive integer } m\}$; that is, $N$ is equivalently the **nilradical** elements of $A$.

Now the projection ring homomorphism

$$\varphi : A \longrightarrow A/N$$

is certainly not an isomorphism (unless $A$ has no nontrivial nilradical elements), but the image of this map under Spec

$$\text{Spec}(\varphi) : \text{Spec}(A/N) \xrightarrow{\sim} \text{Spec}(A)$$

is always an isomorphism. In fact, if we impose the Zarisky topology on these prime spectrums, the functor becomes one which goes to topological spaces

$$\text{Spec}(-) : \textbf{CRing} \longrightarrow \textbf{Top}$$

and the map $\varphi$ becomes a homeomorphism. Hence, this functor does not reflect isomorphisms in either the set or topological senses, because the image $\text{Spec}(\varphi)$ is an isomorphism, but $\varphi$ is not. Despite this, the interpretation of this result is a useful one because it demonstrates that algebraic geometrists can "throw away" their nilradical elements without changing their Zariski topology.

---

**Lemma 1.10.8.** The composition of monomorphisms (epimorphisms) is a (an) monomorphism (epimorphism).

*Proof.* Let $f : A \longrightarrow B$ and $g : B \longrightarrow C$ be monomorphisms, and suppose $h_1, h_2 : D \longrightarrow A$ are two parallel morphisms. Suppose that $(g \circ f) \circ h_1 = (g \circ f) \circ h_2$. Note that we can rewrite the equation to obtain that

$$g \circ (f \circ h_1) = g \circ (g \circ h_1) \implies f \circ h_1 = f \circ h_2.$$

as $g$ is monic, and hence it is left cancellable. But once again, $f$ is monic, so we cancel on the left to obtain that $h_1 = h_2$ as desired. ■

Note: it is not always the case that a monic, epic morphism is an isomorphism (that is, it's not always invertible.)

---

**Example 1.10.9.** Consider the category **Top**, consisting of (small) topological spaces as our objects with continuous functions between them as morphisms. Let $D$ be a dense subset of a topological space $X$ and let $i : D \longrightarrow X$ be the inclusion map. We'll show that this function is both epic and monic.

To show it is epic, let $f_1, f_2 : X \longrightarrow Y$ be continuous maps form $X$ to another topological space $Y$. Let $Y$ be Hausdorff, and suppose that

$$f_1 \circ i = f_2 \circ i.$$

Now $\text{Im}(i) = D$, so the above equation tells us that $f_1(d) = f_2(d)$ for all $d \in D$. That is, the functions agree on the dense subset. However, we know from topology that this implies that $f_1 = f_2$.

*Proof.* Suppose that $f_1(x) \neq f_2(x)$ for some $x \notin D$. Since the points are distinct, and since $Y$ is Hausdorff, there must exist disjoint open sets $U, V$ in $Y$ such that $f_1(x) \in U$ and $f_2(x) \in V$. Since both $f_1, f_2$ are continuous, there must exist open sets $U', V'$ in $X$ such that $f(U') \subseteq U$ and $g(V') \subseteq V$.

However, since $D$ is dense in $X$, both $U'$ and $V'$ must intersect with some portion of $D$; that is, there is some $y \in U'$ and $z \in V'$ such that $y, z \in D$. Therefore, we see that $f_1(y) \in U$ and $f_2(z) \in V$, and since $y, z \in D$ we have that $f_1(y) = f_2(z)$. But this contradicts the fact that $U \cap V = \emptyset$. Therefore, we have a contradiction and it must be the case that $f_1(x) = f_2(x)$ for all $x \in X$, as desired. ■

Therefore, we see that $i$ is epic. To show that it is monic, suppose $g_1, g_2 : Y \longrightarrow D$ are two parallel, continuous functions, and that

$$i \circ g_1 = i \circ g_2.$$

Since $i$ is nothing more than an inclusion map, we immediately have that $g_1 = g_2$. Therefore, $i$ is also monic.

**However**, note that $i : D \longrightarrow X$ is not an isomorphism, since it is not necesasrily always surjective. Hence $i$ is an example of a monic, epic morphism which is not an isomorphism.

---

We finish our discussion on monics and epics by considerig the automorphism groups of a category.

**Definition 1.10.10.** Let $\mathcal{C}$ be a locally small category. For each object $A$ in $\mathcal{C}$, we can consider the **automorphism group** $\text{Aut}(A)$ whose objects consist of isomorphisms $\varphi : A \xrightarrow{\sim} A$, whose product is composition, and whose identity is $1_A$.

Note that despite the notation, this does *not* generally define a functor.

---

**Example 1.10.11.** Some examples of the above construction include familiar and useful examples in mathematics.

- For any group $(G, \cdot)$ in **Grp**, we can formulate the automorphism group $\text{Aut}(G)$ which is the group of isomorphisms from $G$ to itself. Depending on $G$, this can have all kinds of behavior. For example, if $\text{Aut}(G)$ is cyclic, then $G$ is abelian. If $G$ is an abelian group of order $p^n$, then $\text{Aut}(G) = GL(n, F)$ where $F$ is the finite field of order $p$.

- For any set $X$ in **Set**, the automorphism group $\text{Aut}(X)$ consists of the bijections on $X$ to itself; by definition in set theory, these are just permutations. Hence the automorphism group is the permutation group of the elements of $X$.

- For any field $(k, \cdot, +)$ in **Fld**, the automorphism group $\text{Aut}(k)$ also consists of field isomorphisms to itself. In this setting, what is often of more interest is considering the subgroups of $\text{Aut}(k)$, often denoted as $\text{Aut}(k/L)$, which are automorphisms that fix the subfield $L$. These subgroups are key to studying polynomial roots and hence are prevalent in Galois theory.

- For any graph $(G, E, V)$ in **Grph**, one can construct the automorphism group $\text{Aut}(G)$, which tracks the symmetries of the graph. Interestingly, there is a theorem known as Frucht's Theorem which states that every finite group is the automorphism group of a finite (undirected) graph; this was later extended and shown that every group is the automorphism group of a directed graph [*Groups represented by homeomorphism groups.*].

- For any topological space $(X, \tau)$ in **Top**, the autormorphism group $\text{Aut}(X)$ consists of the homeomorphisms to itself. Geometrically, these record the possible ways of continuously deforming a space back into itself. It is a theorem that every group is the automorphism group of some complete, connected, locally connected metric space $M$ of any dimension.

---

With the automorphism group in mind, we might ask the same question on the object level: Given an object $A$ in $\mathcal{C}$, what objects are isomorphic to $A$ in $\mathcal{C}$? To answer this, we define the relation $\sim$ on $\text{Ob}(\mathcal{C})$, the objects of $\mathcal{C}$, where we say

$$A \sim B \text{ if } A \cong B.$$

Such an equivalence relation divides the objects of $\mathcal{C}$ into disjoint *isomorphsm classes*, which reduces the structure of $\mathcal{C}$.

**Definition 1.10.12.** Let $\mathcal{C}$ be a category and $A$ any object. We call the equivalence class of $A$ under $\sim$, defined previously, as the **isomorphism class** which we denote as

$$\text{Isom}(A) = \{X \in \text{Ob}(\mathcal{C}) \mid X \cong A\}.$$

This leads to the following categorical construction which preserves a great deal of information within the category.

**Definition 1.10.13.** Let $\mathcal{C}$ be a category, and assume the axiom of choice. Then we can construct **a skeleton of a category** $\mathcal{C}$, denoted $\text{sk}(\mathcal{C})$, as the category where

**Objects.** For each $A \in \mathcal{C}$, we select one representative of each isomorphism class $\text{Isom}(A)$.

**Morphisms.** For two representatives of isomorphism classes $A, B$, we take

$$\text{Hom}_{\text{sk}(\mathcal{C})}(A, B) = \text{Hom}_{\mathcal{C}}(A, B)$$

We note three things regarding this construction.

**(1)** We used the axiom of choice to build the objects of the category, since we needed to select one element from each isomorphism class.

**(2)** The category $\text{sk}(\mathcal{C})$ is a full subcategory of $\mathcal{C}$ by definition.

**(3)** We note that this construction builds *a* skeleton. In general, a category will have different skeletons because there are many ways to construct the objects of such a skeleton.

As noted, a category will have different skeletons. However, up to isomorphism, it does not really matter which skeleton we build as we will see.

**Lemma 1.10.14.** Let $\mathcal{C}$ be a category, and let $\text{sk}(\mathcal{C})$ and $\text{sk}'(\mathcal{C})$ be two skeletons built from $\mathcal{C}$. Then $\text{sk}(\mathcal{C}) \cong \text{sk}'(\mathcal{C})$.

The prove is left as an exercise for the reader. We will see late that there are more enjoyable properties of "skeletal" categories, which we define as categories exhibiting this type of behavior.

**Definition 1.10.15.** A category $\mathcal{C}$ is called **skeletal** if no two distinct objects are isomorphic in $\mathcal{C}$.

Categorical skeletons are inadvertently studied everywhere in mathematics. For example, asking for a classification of abelian groups, of manifolds, or even of the cardinality of every set is the same thing as asking for the skeletons of **Ab**, **DMan**, and **Set**. We give a few examples.

---

**Example 1.10.16.** Consider the category **FinCard** (read: "finite cardinals") which we describe as

**Objects.** The set $\varnothing$ and the sets $\{1, 2, ..., n\}$ for each $n \in \mathbb{N}$.

**Morphisms.** All functions between these finite sets.

Clearly this is a full subcategory of **FinSet**. Moreover, it is skeletal; no two sets are isomorphic because each object is of different size. Therefore, it is skeletal. In fact, **FinCard** is a skeleton of **FinSet** because any finite set (in some universe $U$) can be ordered in some way, which provides an enumeration on its objects. In other words, every finite set is of some finite size, making it isomorphic to some set $\{0, 1, 2, \dots, n\}$.

---

**Example 1.10.17.** One can try to generalize the previous example to **Set**, but this is in general not possible unless we assume ZFC with the **generalized continuum hypothesis**, as such a posulate is independent of ZFC.

Assuming such an axiom, we can construct the category **Card** where

**Objects.** The sets $\varnothing, \{1, 2, \dots, n\}$ for each $n \in \mathbb{N}$, and $\omega_0, \omega_1, \omega_2, \dots$
**Morphisms.** All functions between such sets.

Here we see that this is again a skeleton **Set**, since by our assumptions (which is assuming a lot), any set is of some cardinality $1, 2, \dots, n, \dots, \aleph_0, \aleph_1, \dots$. However, for each such cardinal we have a corresponding set with that cardinality. Hence each element in **Set** is isomorphic to some element of **Card**. Overall, we see that **Card** forms a skeleton of **Set**.

---

The above example can be repeated for **Cycl**, the category of cyclic groups. This is because any two cyclic groups of the same order are isomorphic. Hence, one can find a skeleton of **Cycl** by finding a family of cylic groups of every set size (again, using the generalized continuum hypothesis).

---

**Example 1.10.18.** Consider the category **Ecld** of Euclidean spaces, which we may describe as

**Objects.** The vector spaces $\mathbb{R}^n$ for each $n = 0, 1, 2, \dots,$
**Morphisms.** Linear transformations between vector spaces.

Then we see that **Ecld** is the skeleton of $\mathbf{FinVect}_k$, which is the category of finite-dimensional vector spaces. The reason why this works is because every finite dimensional vector space is isomorphic to $\mathbb{R}^n$ for some $n$.

---

## Exercises

1. Prove Lemma 1.10.8 for epimorphisms.
2. Prove Lemma 1.10.14.

**3.** Describe the monomorphisms and epimorphisms in the category of **Cat**.[10]

**4.** In the category of **Ring**, give an example of a morphism which is both a monomorphism and epimorphism, but not an isomorphism.
(*Hint:* Consider the inclusion $i : \mathbb{Z} \longrightarrow \mathbb{Q}$.)

**5.** Recall from Exercise ? that, in any category, if we have two commutative diagrams, we can always stack them together to obtain a larger commutative diagram. We saw, however, that converse is not always true: subdividing a commutative diagram does not produce smaller commutative diagrams.

Prove that the converse is true when all morphisms are isomorphisms.

[10]Classifying epimorphisms in **Cat** is actually nontrivial, although not impossible. However, the task here is to just interpret the definition of monics and epics **Cat**.

## 1.11 Initial, Terminal, and Zero Objects

We can also be more specific in discussing the nature of the objects of a given category $\mathcal{C}$.

**Definition 1.11.1.** Let the following objects exist in some category $\mathcal{C}$.

- Let $T$ be an object. Then $T$ is **terminal** if for each object $A$, there exists exactly one morphism $f_A$ such that $f_A : A \longrightarrow T$.
- Let $I$ be an object. Then $I$ is said to be **initial** if for each object $A$ there exists exactly one morphism $f_A : I \longrightarrow A$.
- An object $Z$ is said to be a **zero object** if it is both terminal and initial. Since terminal and initial objects are unique, so is a zero object.

  Equivalently, it is zero if for any objects $A, B$, there exists exactly one morphism $f : A \longrightarrow Z$ and exactly one morphism $g : Z \longrightarrow B$. Hence, for any two objects there exists a morphism between them, namely given by by $g \circ f$, called the **zero morphism** from $A$ to $B$.

If an object $T$ is terminal, then there is one and only morphism to itself (namely, its identity). Therefore, for any two terminal objects $T$ and $T'$, they are isomorphic, since by assumption there exists unique morphisms $f : T \longrightarrow T'$ and $g : T' \longrightarrow T$ and we have no choice but to say

$$f \circ g = 1_T \quad g \circ f = 1_{T'}.$$

---

**Example 1.11.2.** Recall that in the category **Grp**, there exists a trivial group $\{e\}$. Moreover, for each group $G$, there exist unique group homomorphisms

$$i_G : \{e\} \longrightarrow G \qquad e \mapsto e_G$$

and

$$t_G : G \longrightarrow \{e\} \qquad g \mapsto e_G.$$

Note that both are group homomorphisms since they both behave on identity elements and are trivially distributive across group operations. This then shows that **Grp**, the trivial group is initial and terminal and hence a zero object.

This makes sense since for any two groups $G, H$, there exists a unique map

$$z : G \longrightarrow H \qquad g \mapsto e_H$$

which could be factorized as

$$\begin{tikzcd} & \{e\} \arrow[dr, "i_H"] & \\ G \arrow[ur, "t_G"] \arrow[rr, "z"'] & & H \end{tikzcd}$$

which demonstrates the existence of a zero object (the name "zero" makes sense now, right?), which we already know is $\{e\}$. Note in this example, we did not actually use much group theory. In fact, this could be repeated for the categories $R$ mod , **Ab**, and other similar categories.

---

The next two examples demonstrate that terminal and initial objects of course don't always have to coincide like they did in the previous example.

---

**Example 1.11.3.** Let $n$ be a positive integer. Recall that we can create a category, specifically a preorder, by taking our objects to be positive integers less than $n$, and allowing one morphism $f : k \longrightarrow m$ whenever $k \leq m$.

$$\begin{tikzcd} 1 \arrow[r] \arrow[loop above] & 2 \arrow[r] \arrow[loop above] & 3 \arrow[r] \arrow[loop above] & \cdots \arrow[r] & n \arrow[loop above] \end{tikzcd}$$

Then 1 is an initial object while $n$ is a terminal object. This is because for any number $1 \leq m \leq n$, there exists a unique morphism from 1 to $m$, and a unique morphism $m$ to $n$, both which may be obtained by repeated composition.

---

---

**Example 1.11.4.** Consider the category **Set**. Let $X$ be a given set in this category. Then there are two unique functions which we may construct. First, there is the function

$$t_X : X \longrightarrow \{\bullet\}$$

where everything in $X$ is mapped to the one element $\bullet$ of the one point set. Secondly, we may construct a function whose domain is the empty set, and whose codomain is $X$, as below.

$$i_X : \varnothing \longrightarrow X$$

Thus we have that, in **Set**, the one point set is a terminal object $\{\bullet\}$ while the empty set $\varnothing$ is an initial object.

One may wonder at this point: How exactly is $i_X$ a true, set theoretic function? And why can't we also obtain a unique morphism $i'_X : X \longrightarrow \varnothing$, so that $\varnothing$ is a terminal object as well?

The second question is easy to answer; if $\varnothing$ was also terminal, then we'd have that $\{\bullet\} \cong \varnothing$ which is not true. Since this is a bit of a boring answer, we'll explain in detail.

Recall that a function in $f : A \longrightarrow X$ between two sets $A$ and $X$ is a relation $R \subseteq A \times X$ which satisfies two properties.

1. (Existence.) For each $a \in A$, there exists a $x \in X$ such that $(a, x) \in R$
2. (Uniqueness. Or, if you'd like, the vertical line test.) If $(a, x) \in R$ and $(a, x') \in R$ then $x = x'$.

Now observe that if $A = \varnothing$, then $R \subseteq \varnothing \times X = \varnothing$. Hence (1) and (2) are satisfied because each is trivially true. However, we don't get a function $f : X \longrightarrow \varnothing$, since in this case (1) fails. Specifically, (1) demands the existence of elements in our codomain, a demand we cannot meet if it is empty.

Thus we see that $\varnothing$ is initial, but not terminal as our intuition may suggest, and that $\{\bullet\}$ is terminal.

---

**Example 1.11.5.** Consider the category of fields **Fld**. Suppose we ask if this has an initial or terminal object.

We might guess that the smallest field

$$\mathbb{F}_2 \cong (\mathbb{Z}/2\mathbb{Z}, +, \cdot) = \{0, 1\}$$

which has characteristic 2 is an initial object. However, this fails to be initial. Observe that the only homomorphism between $\mathbb{F}_2$ and $\mathbb{F}_3$ is the zero homomorphism, which is not in our category. (Recall that **Fld** is a full subcategory of **Ring**, a category whose morphisms we require to be unit preserving.)

The reason why it must be the zero homomorphisms is because $\mathbb{F}_3$ has characteristic three, and in general, two fields will only share a (nonzero) field homomorphisms if they have the same characteristic.

By a similar argument, we can state that terminal objects also do no exist. Overall, these objects fail to exist in **Fld** because fields have a large set of restictions imposed by their numerous axioms. Hence, this category lacks initial and terminal objects.

---

## Exercises

**1.** $(i.)$ Let $\mathcal{C}$ be a category with initial object $I$. For any two objects $A, B \in \mathcal{C}$, define for each $f \in \text{Hom}_{\mathcal{C}}(A, B)$ the functor

$$P_f : \mathbf{2} \longrightarrow \mathcal{C}$$

such that $P(\bullet) = A$, $P(\bullet) = B$, and $P_f(\bullet \longrightarrow \bullet) = f : A \longrightarrow B$. Show that for each

$f : A \longrightarrow B$ in $\mathcal{C}$, we have a natural transformation

$$\eta : P_{1_I} \longrightarrow P_f.$$

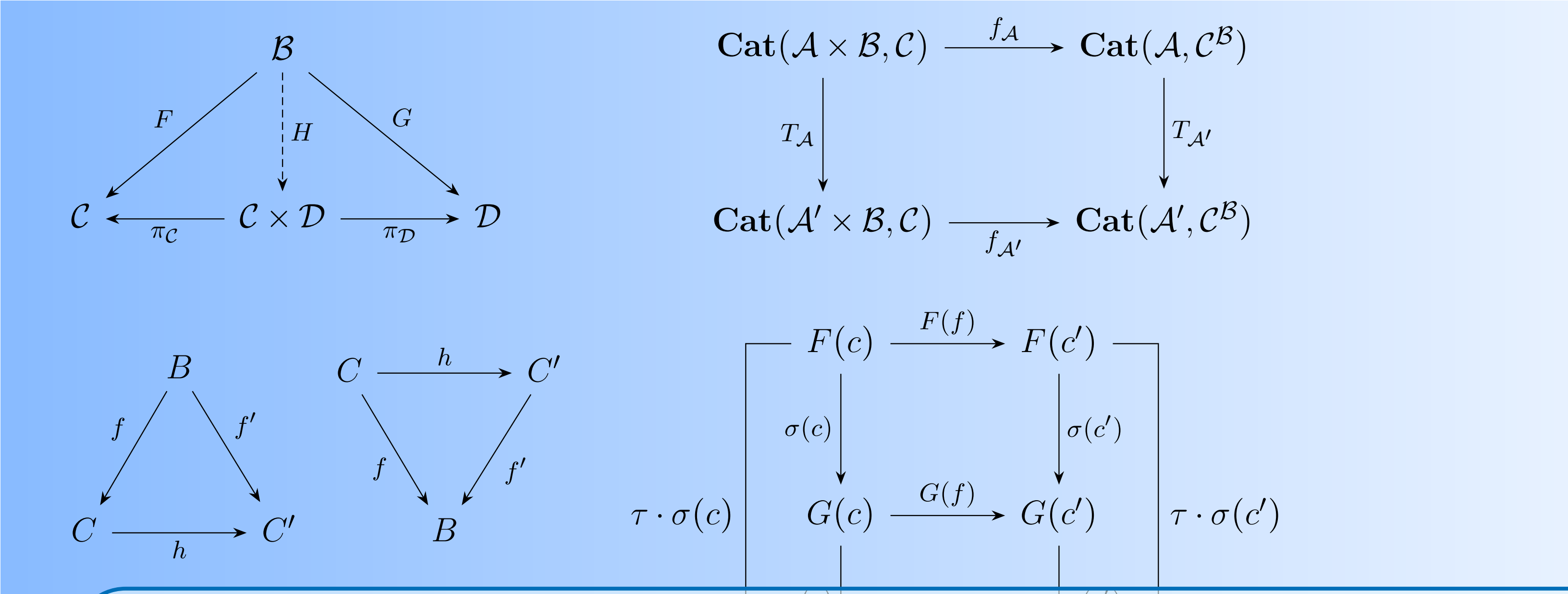
Note that $1_I : I \longrightarrow I$ is the identity on the initial object.

(*ii.*) Suppose we don't know if $\mathcal{C}$ has an initial object, but we have a distinguished object $I'$ with the property that for each $f \in \mathrm{Hom}_{\mathcal{C}}(A, B)$ there is a natural transformation

$$\eta : P_{1_{I'}} \longrightarrow P_f.$$

Is $I'$ an inital object?

(*iii.*) Dualize your work for terminal objects.
(*Hint*: We now want a natural transformation $\eta' : P_f \longrightarrow P_{1_I}$).



# 2. Duality and Categorical Constructions

## 2.1 $\mathcal{C}^{\text{op}}$ and Contravariance

**Definition 2.1.1.** Consider a category $\mathcal{C}$. Then we define the **opposite category of** $\mathcal{C}$, denoted $\mathcal{C}^{\text{op}}$, to be the category where

**Objects.** The same objects of $\mathcal{C}$.

**Morphisms.** If $f : A \longrightarrow B$ is a morphism of $\mathcal{C}$, then we let $f^{\text{op}} : B \longrightarrow A$ be a morphism of $\mathcal{C}^{\text{op}}$.

In this case, composition isn't exactly obvious, so we will explain how that works.

Let $f : A \longrightarrow B$ and $g : B \longrightarrow C$ be morphisms of $\mathcal{C}$. Then we obtain morphisms $f^{\text{op}} : B \longrightarrow A$ and $g^{\text{op}} : C \longrightarrow B$. In this case $f^{\text{op}}, g^{\text{op}}$ are composable, and we define composition of $\mathcal{C}^{\text{op}}$, denoted as $\circ^{\text{op}}$, to be the morphism

$$f^{\text{op}} \circ g^{\text{op}} : C \longrightarrow A.$$

Moreover, we have the relation $(g \circ f)^{\text{op}} = f^{\text{op}} \circ g^{\text{op}}$.

Taking the opposite category might seem very strange, but we are doing nothing more than just taking the same category and swapping the domain and codomain of every morphism.

Consequently, many properties of morphisms are similarly reversed. For example, if $f : A \longrightarrow B$ is monomorphism in $\mathcal{C}$, then $f^{\text{op}} : B \longrightarrow A$ is an epimorphism in $\mathcal{C}^{\text{op}}$. More generally, every logically valid statement that can be made in $\mathcal{C}$ using its objects and morphisms can be dualized to achieve an equivalent, logically valid statement in $\mathcal{C}^{\text{op}}$ using its objects and morphisms.

---

**Example 2.1.2.** Consider a category $\mathcal{C}$ containing 3 objects whose morphisms are arranged as follows:

$$
\begin{tikzcd}
 & A \arrow[loop above, "1_A"] \arrow[dr, "f"] & \\
C \arrow[loop above, "1_C"] \arrow[ur, "h"] & & B \arrow[loop above, "1_B"] \arrow[ll, "g"]
\end{tikzcd}
$$

What does the dual category $\mathcal{C}^{\text{op}}$ look like? Well, $\mathcal{C}^{\text{op}}$ contains the same objects $A, B$ and $C$. As for the morphisms, $\mathcal{C}$ has the three morphisms $f, g, h$, in addition to their composites. Therefore, $\mathcal{C}^{\text{op}}$ also has three morphisms $f^{\text{op}} : B \longrightarrow A$, $g^{\text{op}} : C \longrightarrow B$ and $h^{\text{op}} : A \longrightarrow C$ and their composites. Hence, $\mathcal{C}^{\text{op}}$ looks like this:

$$
\begin{tikzcd}
 & A \arrow[loop above, "1_A"] \arrow[dl, "h"'] & \\
C \arrow[loop above, "1_C"] \arrow[rr, "g"'] & & B \arrow[loop above, "1_B"] \arrow[ul, "f"']
\end{tikzcd}
$$

---

---

**Example 2.1.3.** Let $P$ be a preorder, specifically a partial order. Recall that this means that $P$ has a binary relation $\le$ and if $p \le p'$ and $p' \le p$, then $p = p'$.

We claim that that $P^{\text{op}}$ is still a partial order. But first, what does $P^{\text{op}}$ even look like? If we have some elements $p_1, p_2, p_3$ in $P$ such that

$$p_1 \le p_2 \le p_3$$

Then, as a category, $P$ has the unique morphisms $f : p_1 \longrightarrow p_2$ and $g : p_2 \longrightarrow p_3$. Hence, in $P^{\text{op}}$, we have the unique morphisms $g^{\text{op}} : p_3 \longrightarrow p_2$ and $f^{\text{op}} : p_2 \longrightarrow p_1$, so that we obtain a reversed binary relation $\le^{\text{op}}$ in $P$, which reorder $p_1, p_2, p_3$ as below.

$$p_3 \le^{\text{op}} p_2 \le^{\text{op}} p_1$$

This is kinda weird to write, and in fact, it makes more sense if we write $\le^{\text{op}} = \ge$ as the binary relation in $P^{\text{op}}$. We then have that

$$p_1 \le p_2 \le p_3 \text{ in } P \implies p_3 \ge p_2 \ge p_1 \text{ in } P^{\text{op}}$$

which is nice! Things are even nicer in a linear order, for if $P = \{p_1, p_2, p_3, \dots\}$ is a linear order, then we can write that

$$\cdots p_i \le p_j \le p_k \cdots$$

and hence in $P^{\text{op}}$ this becomes

$$\cdots p_i \ge p_j \ge p_k \cdots .$$

**Example 2.1.4.** Let $(G, \cdot)$ be a group. In group theory one can formulate the **opposite group** $(G^{\text{op}}, \cdot^{\text{op}})$ as follows. Define $(G^{\text{op}}, \cdot^{\text{op}})$ to be group with the same set of elements as $G$, whose product $\cdot^{\text{op}}$ works as

$$g_1 \cdot^{\text{op}} g_2 = g_2 \cdot g_1.$$

Since both $(G, \cdot)$ and $(G, \cdot^{\text{op}})$ are groups, we can regard them both as one object categories. What is interesting to realize is that under the categorical interpretation, they are opposite categories of each other.

We thus see that dualizing a category simply involves changing the directions of the morphisms on the objects. But can we dualize a functor?

**Definition 2.1.5.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor and suppose $f : A \longrightarrow B$ is morphism in $\mathcal{C}$. We say $F$ is a **contravariant functor** if $F(f) : F(B) \longrightarrow F(A)$.

This is in sharp contrast to a *covariant* functor, in which $f : A \longrightarrow B$ is sent to $F(f) : F(A) \longrightarrow F(B)$.

We next introduce a few examples to demonstrate a contravariant functor.

**Example 2.1.6.** Let $k$ be an algebraically closed field. Recall that $A^n(k)$ is the set of tuples $(a_1, a_2, \dots, a_n)$ with $a_i \in k$. In algebraic geometry, it is of interest to associate each subset $S \subseteq A^n(k)$ with the ideal

$$I(S) = \left\{ f \in k[x_1, \dots, k_n] \;\middle|\; f(s) = 0 \text{ for all } s \in S \right\}.$$

of $k[x_1, \dots, x_n]$. Observe that this is always non-empty since $0 \in I(S)$ for any $S$. In additional, it is clearly an ideal of $k[x_1, \dots, x_n]$, since for any $p \in k[x_1, \dots, x_n]$,$q \in I(S)$, we have that

$$(p \cdot q)(s) = p(s) \cdot q(s) = p(s) \cdot 0 = 0 \text{ for all } s \in S.$$

so that $p \cdot q \in I(S)$. Now it's usually an exercise to show that if $S_1 \subseteq S_2$ are two subsets of $A^n(k)$, then one has that $I(S_2) \subseteq I(S_1)$. Hence this defines a contravariant functor

$$I : \textbf{Subsets}(A^n(k)) \longrightarrow \textbf{Ideals}(k[x_1, \dots, x_n]).$$

where $\textbf{Subsets}(A^n(k))$ is the category of subsets with inclusion morphisms, and $\textbf{Ideals}(k[x_1, \dots, x_n])$ is the category of ideals with inclusion ring homomorphisms.

**Example 2.1.7.** Consider again $k$ as an algebraically closed field. In algebraic geometry, one often wishes to associated each ideal of $k[x_1, \dots, x_n]$ with its "zero set"

$$Z(I) = \left\{ s = (a_1, \dots, a_n) \in A^n(k) \;\middle|\; f(s) = 0 \text{ for all } s \in I \right\}.$$

It is usually an exercise to show that if $I_1 \subseteq I_2$ are two ideals, then $Z(I_2) \subseteq Z(I_1)$. Hence we see that this defines a contravariant functor

$$Z : \textbf{Ideals}(k[x_1, \dots, x_n]) \longrightarrow \textbf{Subsets}(A^n(k)).$$

---

It is usually at the beginning of an algebraic geometry course that one will understand the relationship between these two constructions, which themselves are secretly functors.

What follows is a very interesting example. In fact, this example is an example of a beautiful concept of a *sheaf*, and it is usually used as a motivating example. But that is for later.

---

**Example 2.1.8.** Let $X$ be a topological space, and consider the thin category $\textbf{Open}(X)$, which contains all open sets $U \subseteq X$, equipped with the inclusion function $i_{U,X} : U \longrightarrow X$.

For each $U \in \textbf{Open}(X)$, define the set

$$C(U) = \{f : U \longrightarrow \mathbb{R} \mid f \text{ is continuous.}\}$$

Note that if $U \subseteq V$ are in $\textbf{Open}(X)$, then we define the function $\rho_{U,V} : C(V) \longrightarrow C(U)$ where

$$\rho_{U,V}(f : V \longrightarrow \mathbb{R}) = f\Big|_U : U \longrightarrow \mathbb{R}.$$

That is, $\rho_{U,V}$ sends continuous, real-valued functions on $V$ to such functions on $U$ by restriction. It is not difficult to show that this respects identity and composition requirements, so that we have a contravariant functor

$$C(-) : \textbf{Open}(X) \longrightarrow \textbf{Set}$$

for each topological space $X$.

---

What follows is another very important example.

---

**Example 2.1.9.** Let $\mathcal{C}$ be a locally small category. In this case, we know that each $A \in \mathcal{C}$ induces the covariant functor

$$\text{Hom}_{\mathcal{C}}(A, -) : \mathcal{C} \longrightarrow \textbf{Set}$$

which sends objects $C$ to the set $\text{Hom}_{\mathcal{C}}(A, C)$. It is natural to ask if we may similarly define a functor

$$\text{Hom}_{\mathcal{C}}(-, A) : \mathcal{C} \longrightarrow \textbf{Set}.$$

The answer is yes. We did not make this observation in the past for pedagogical reasons, since it's actually a contravariant functor (and we didn't know what that was until now). We can now safely say that $\text{Hom}_{\mathcal{C}}(-, A)$ is a contravariant functor.

---

We now comment on the relationship between contravariant and covariant functors.

**Proposition 2.1.10.** Let $\mathcal{C}$, $\mathcal{D}$ be categories.

- Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a contravariant functor. Then $F$ corresponds to a contravariant functor $\overline{F} : \mathcal{C}^{\text{op}} \longrightarrow \mathcal{D}$ where for a $f^{\text{op}} : B \longrightarrow A \in \mathcal{C}^{\text{op}}$,

$$\overline{F}(f^{\text{op}} : B \longrightarrow A) = F(f : A \longrightarrow B) = F(f) : F(B) \longrightarrow F(A).$$

- Conversely, let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a covariant functor. Then $F$ corresponds to a contravariant functor $\overline{F} : \mathcal{C}^{\text{op}} \longrightarrow \mathcal{D}$ where

$$\overline{F}(f^{\text{op}} : B \longrightarrow A) = F(f : A \longrightarrow B) = F(f) : F(A) \longrightarrow F(B)$$

The above proposition allows us to treat any functor as covariant or contravariant. Thus, if we don't like the behavior of our functor on morphisms, we can find an equivalent functor that behaves on morphisms in our preferred way.

Generally, covariant functors are easier to think about, so we often like to turn contravariant functors into covariant functors.

---

**Example 2.1.11.** Recall that the functor

$$C(-) : \textbf{Open}(X) \longrightarrow \textbf{Set}$$

is contravariant. What if we want to treat this as a covariant functor? Well, we can define the functor

$$\overline{C}(-) : \textbf{Open}(X)^{\text{op}} \longrightarrow \textbf{Set}$$

as follows. If $U \subseteq V$ are open subsets of the topological space $X$, then let $i : U \longrightarrow V$ be the inclusion. This is a morphism in $\textbf{Open}(X)$. Hence, $i^{\text{op}} : V \longrightarrow U$ is a morphism in $\textbf{Open}(X)^{\text{op}}$. Therefore, we define

$$\overline{C}(i^{\text{op}} : V \longrightarrow U) = C(i : U \longrightarrow V) = \rho_{U,V} : C(V) \longrightarrow C(U).$$

Thus we see that this functor $\overline{C}$ acts the same way as $C$, except it behaves covariantly on the morphisms now instead of contravariantly.

---

## 2.2 Products of Categories, Functors

As one may expect, the product of categories can be easily defined.

**Definition 2.2.1.** Let $\mathcal{C}$ and $\mathcal{D}$ be categories. Then the **product category** $\mathcal{C} \times \mathcal{D}$ is the category where
**Objects.** All pairs $(C, D)$ with $C \in \text{Ob}(\mathcal{C})$ and $D \in \text{Ob}(\mathcal{D})$
**Morphisms.** All pairs $(f, g)$ where $f \in \text{Hom}(\mathcal{C})$ and $g \in \text{Hom}(\mathcal{D})$.
To define composition in this category, suppose we have composable morphisms in $\mathcal{C}$ and $\mathcal{D}$ as below.

$$\mathcal{C} \qquad \begin{tikzcd} \cdots \; C_1 \arrow[r, "f"] \arrow[rr, bend left, "f' \circ f"] & C_2 \arrow[r, "f'"] & C_3 \; \cdots \end{tikzcd}$$

$$\mathcal{D} \qquad \begin{tikzcd} \cdots \; D_1 \arrow[r, "g"] \arrow[rr, bend left, "g' \circ g"] & D_2 \arrow[r, "g'"] & D_3 \; \cdots \end{tikzcd}$$

Then the morphisms $(f, g)$ and $(f', g')$ in $\mathcal{C} \times \mathcal{D}$ are composable too, and their composition is defined as $(f', g') \circ (f, g) = (f' \circ f, g' \circ g)$.

$$\mathcal{C} \times \mathcal{D} \qquad \begin{tikzcd} \cdots \; (C_1, D_1) \arrow[r, "(f,g)"] \arrow[rr, bend left, "(f',g') \circ (f,g) = (f' \circ f, g' \circ g)"] & (C_2, D_2) \arrow[r, "(f',g')"] & (C_3, D_3) \; \cdots \end{tikzcd}$$

We also define the **projection functors** $\pi_{\mathcal{C}} : \mathcal{C} \times \mathcal{D} \longrightarrow \mathcal{C}$ and $\pi_{\mathcal{D}} : \mathcal{C} \times \mathcal{D} \longrightarrow \mathcal{D}$ where on objects $(C, D)$ and morphism $(f, g)$, we have that

$$\pi_{\mathcal{C}}(C, D) = C \qquad\qquad \pi_{\mathcal{D}}(C, D) = D$$
$$\pi_{\mathcal{C}}(f, g) = f \qquad\qquad \pi_{\mathcal{D}}(f, g) = g$$

These projection functors have the following property. Consider a pair of functors $F : \mathcal{B} \longrightarrow \mathcal{C}$ and $G : \mathcal{B} \longrightarrow \mathcal{D}$. Then $F$ and $G$ determine a unique functor $H : \mathcal{B} \longrightarrow \mathcal{C} \times \mathcal{D}$ where

$$\pi_{\mathcal{C}} \circ H = F \qquad \pi_{\mathcal{D}} \circ H = G.$$

That is, we see that for any morphism $f$ in $\mathcal{B}$ we have that $H(f) = (F(f), G(f))$. Hence the following diagram commutes

$$\begin{tikzcd} & \mathcal{B} \arrow[dl, "F"'] \arrow[d, dashed, "H"] \arrow[dr, "G"] & \\ \mathcal{C} & \mathcal{C} \times \mathcal{D} \arrow[l, "\pi_{\mathcal{C}}"] \arrow[r, "\pi_{\mathcal{D}}"'] & \mathcal{D} \end{tikzcd}$$

and we dash the middle arrow to represent that $H$ is induced, or defined, by this process.

We can also take the product of two different functors.

**Definition 2.2.2.** Let $F : \mathcal{C} \longrightarrow \mathcal{C}'$ and $G : \mathcal{D} \longrightarrow \mathcal{D}'$ be two functors. Then we define the **product functor** to be the functor $F \times G : \mathcal{C} \times \mathcal{D} \longrightarrow \mathcal{C}' \times \mathcal{D}'$ for which

1. If $(C, D)$ is an object of $\mathcal{C} \times \mathcal{D}$ then $(F \times G)(C, D) = (F(C), G(D))$
2. If $(f, g)$ is a morphism of $\mathcal{C} \times \mathcal{D}$ then $(F \times G)(f, g) = (F(f), G(g))$

   Additionally, we can compose the product of functors (of course, so long as they have the same number of factors). Thus suppose $G, F$ and $G', F'$ are composable functors. Then observe that
$$(G \times G') \circ (F \times F') = (G \circ F) \times (G' \circ F').$$

Note that in this formulation we have that

$$\pi_{\mathcal{C}'} \circ (F \times G) = F \circ \pi_{\mathcal{C}} \quad \pi_{\mathcal{C}'} \circ (F \times G) = G \circ \pi_{\mathcal{D}}$$

Hence, we have the following commutative diagram.

$$\begin{tikzcd}
\mathcal{C} \arrow[d, "F"'] & \mathcal{C} \times \mathcal{D} \arrow[l, "\pi_{\mathcal{C}}"'] \arrow[r, "\pi_{\mathcal{D}}"] \arrow[d, dashed, "F \times G"] & \mathcal{D} \arrow[d, "G"] \\
\mathcal{C}' & \mathcal{C}' \times \mathcal{D}' \arrow[l, "\pi_{\mathcal{C}'}"'] \arrow[r, "\pi_{\mathcal{D}'}"] & \mathcal{D}'
\end{tikzcd}$$

Again, the dashed arrow is written to express that $F \times G$ is the functor defined by this process and makes this diagram commutative.

**Definition 2.2.3.** If $F$ is a functor such that $F : \mathcal{B} \times \mathcal{C} \longrightarrow \mathcal{D}$, that is, its domain is a product category, then $F$ is said to be a **bifunctor**.

An example of a bifunctor is the cartesian product $\times$, which we can apply to sets, groups, and topological spaces. In these instances we know that value of a cartesian product is always determined uniquely by the values of the individual factors, which holds more generally for bifunctors.

**Proposition 2.2.4.** Let $\mathcal{B}, \mathcal{C}$ and $\mathcal{D}$ be categories. For $B \in \mathcal{B}$ and $C \in \mathcal{C}$, define the functors

$$H_C : \mathcal{B} \longrightarrow \mathcal{D} \quad K_B : \mathcal{C} \longrightarrow \mathcal{D}$$

such that $H_C(B) = K_B(C)$ for all $B, C$. Then there exists a functor $F : \mathcal{B} \times \mathcal{C} \longrightarrow \mathcal{D}$ where $F(B, -) = K_B$ and $F(-, C) = H_C$ for all $B, C$ if and only if for every pair of morphisms $f : B \longrightarrow B'$ and $g : C \longrightarrow C'$ we have that

$$K_{B'}(g) \circ H_C(f) = H_{C'}(f) \circ K_B(g).$$

Diagrammatically, this condition is

$$
\begin{tikzcd}
H_C(B) = K_B(C) \arrow[r, "K_B(g)"] \arrow[d, "H_C(f)"'] & H_{C'}(B) = K_B(C') \arrow[d, "H_{C'}(f)"] \\
H_C(B') = K_{B'}(g) \arrow[r, "K_{B'}(g)"'] & H_{C'}(B') = K_{B'}(C')
\end{tikzcd}
$$

The proof is left as an exercise for the reader.

---

**Example 2.2.5.** We now introduce what is probably one of the most important examples of a bifunctor. Note that for any (locally small) category $\mathcal{C}$, we have for each object $A$ a functor.

$$\text{Hom}(A, -) : \mathcal{C} \longrightarrow \textbf{Set}$$

We also have a functor from $\mathcal{C}^{\text{op}}$ (we at the $^{\text{op}}$ simply for convenience) for each $B \in \mathcal{C}^{\text{op}}$.

$$\text{Hom}(-, B) : \mathcal{C}^{\text{op}} \longrightarrow \textbf{Set}$$

As an application of the proposition, one can see that that these two functors act as the $K_B$ and $H_C$ functors in the above proposition, and give rise to bifunctor

$$\text{Hom} : \mathcal{C}^{\text{op}} \times \mathcal{C} \longrightarrow \textbf{Set}.$$

This is because for any $h : A \longrightarrow A'$ and $k : B \longrightarrow B'$, the diagram,

$$
\begin{tikzcd}
\text{Hom}(A', B) \arrow[r, "h^*"] \arrow[d, "k_*"'] & \text{Hom}(A, B) \arrow[d, "k_*"] \\
\text{Hom}(A', B') \arrow[r, "h^*"] & \text{Hom}(A, B')
\end{tikzcd}
$$

commutes. Hence the proposition guarantees that $\text{Hom} : \mathcal{C}^{\text{op}} \times \mathcal{C} \longrightarrow \textbf{Set}$ exists and is unique.

---

---

**Example 2.2.6.** Recall that for an integer $n$ and for a ring $R$ with identity $1 \neq 0$, we can formulate the group $\text{GL}(n, R)$, consisting of $n \times n$ matrices with entry values in $R$. As this takes in arguments, we might guess that we have a bifunctor

$$GL(-, -) : \mathbb{N} \times \textbf{Ring} \longrightarrow \textbf{Grp}$$

where $\mathbb{N}$ is a the discrete category with elements as natural numbers. This intuition is correct: for a fixed ring $R$, we have a functor

$$GL(-, R) : \mathbb{N} \longrightarrow \textbf{Grp}$$

while for a fixed natural number $n$ we have a functor

$$GL(n, -) : \mathbf{Ring} \longrightarrow \mathbf{Grp}.$$

Below we can visualize the activity of this functor:

$$\begin{array}{c|ccccc}
\vdots & \vdots & \vdots & \cdots & \vdots & \cdots \\
R = S & GL(1,S) & GL(2,S) & \cdots & GL(k,S) & \cdots \\
\vdots & \vdots & \vdots & \cdots & \vdots & \cdots \\
R = \mathbb{Z} & GL(1,\mathbb{Z}) & GL(2,\mathbb{Z}) & \cdots & GL(k,\mathbb{Z}) & \cdots \\
\hline
 & n=1 & n=2 & \cdots & n=k & \cdots
\end{array}$$

Above, we start with $\mathbb{Z}$ since the this is the initial object of the category **Ring**.

---

Now that we understand products of categories a functors, and we have a necessary and sufficient condition for the existence of a bifunctor, we describe necessary and sufficient conditions for the existence of a natural transformation.

**Definition 2.2.7.** Suppose $F, G : \mathcal{B} \times \mathcal{C} \longrightarrow \mathcal{D}$ are bifunctors. Suppose that there exists a morphism $\eta$ which assigns objects of $\mathcal{B} \times \mathcal{C}$ to morphisms of $\mathcal{D}$. Specifically, $\eta$ assigns objects $B \in \mathcal{B}$ and $C \in \mathcal{C}$ to the morphism

$$\eta_{(B,C)} : F(B,C) \longrightarrow G(B,C).$$

Then $\eta$ is said to be **natural** in $B$ if, for all $C \in \mathcal{C}$,

$$\eta_{(-,C)} : F(-,C) \longrightarrow G(-,C)$$

is a natural transformation of functors from $\mathcal{B} \longrightarrow \mathcal{D}$.

With the previous definition, we can now introduce the necessary condition for a natural transformation to exist between bifunctors.

**Proposition 2.2.8.** Let $F, G : \mathcal{B} \times \mathcal{C} \longrightarrow \mathcal{D}$ be bifunctors. Then there exists a natural transformation $\eta : F \longrightarrow G$ if and only if $\eta(B,C)$ is natural in $B$ for each $C \in C$, and natural in $C$ for each $B \in \mathcal{B}$.

*Proof.*
( $\Longrightarrow$ ) Suppose that $\eta : F \longrightarrow G$ is a natural transformation. Then every object $(B,C)$ is associated with a morphism $\eta_{(B,C)} : F(B,C) \longrightarrow G(B,C)$ in $\mathcal{D}$, and this gives rise to the following diagram:

$$
\begin{tikzcd}
(B,C) \arrow[d, "{(f,g)}"] \\
(B',C')
\end{tikzcd}
\qquad
\begin{tikzcd}
F(B,C) \arrow[r, "\eta_{(B,C)}"] \arrow[d, "{F(f,g)}"'] & G(B,C) \arrow[d, "{G(f,g)}"] \\
F(B',C') \arrow[r, "\eta_{(B',C')}"] & G(B',C')
\end{tikzcd}
$$

Now let $C \in \mathcal{C}$ and observe that

$$\eta_{(-,C)} : F(-,C) \longrightarrow G(-,c)$$

is a natural transformation for all $B$. On the other hand, for any $B \in \mathcal{B}$,

$$\eta_{(B,-)} : F(B,-) \longrightarrow G(B,-)$$

is a natural transformation for all $C$. Therefore, $\eta$ is both natural in $B$ and $C$ for all objects $(B,C)$

( $\Longleftarrow$ ) Suppose on the other hand that $\eta$ is a function which assigns objects $(B,C)$ to a morphism $F(B,C) \longrightarrow G(B,C)$ in $\mathcal{D}$. Furthermore, suppose that $\eta(B,C)$ is natural in $B$ for all $C \in \mathcal{C}$ and natural in $C$ for all $B \in \mathcal{B}$.

Consider a morphism $(f,g) : (B,C) \longrightarrow (B',C')$ in $\mathcal{B} \times \mathcal{C}$. Then since $\eta$ is natural for all $B \in \mathcal{B}$, we know that for all $C \in \mathcal{C}$,

$$\eta_{(-,C)} : F(-,C) \longrightarrow G(-,C)$$

is a natural transformation. In addition, $\eta$ is natural for all $C \in \mathcal{C}$ since for all $B \in \mathcal{B}$

$$\eta_{(B,-)} : F(B,-) \longrightarrow G(B,-)$$

is a natural transformation. Hence consider the natural transformation $\eta_{(-,C)}$ acting on $(B,C)$ and $\eta_{(B',-)}$ acting on $(B',C)$. Then we get the following commutative diagrams.

$$
\begin{tikzcd}
F(B,C) \arrow[r, "\eta_{(B,C)}"] \arrow[d, "{F(f,1_C)}"'] & G(B,C) \arrow[d, "{G(f,1_C)}"] \\
F(B',C) \arrow[r, "\eta_{(B',C)}"] & G(B',C)
\end{tikzcd}
\qquad
\begin{tikzcd}
F(B',C) \arrow[r, "\eta_{(B',C)}"] \arrow[d, "{F(1_{B'},g)}"'] & G(B',C) \arrow[d, "{G(1_{B'},g)}"] \\
F(B',C') \arrow[r, "\eta_{(B',C')}"] & G(B',C')
\end{tikzcd}
$$

Observe that the bottom row of the first diagram matches the top row of the second. Also note that $f : B \longrightarrow B'$ and $g : C \longrightarrow C'$, and that the diagrams imply the equations

$$G(f,1_C) \circ \eta_{(B,C)} = \eta_{(B',C)} \circ F(f,1_C) \tag{2.1}$$
$$G(1_{B'},g) \circ \eta_{(B',C)} = \eta_{(B',C')} \circ F(1_{B'},g). \tag{2.2}$$

Now suppose we compose equation (2.1) with $G(1_{B'}, g)$ on the left. Then we get that

$$\begin{aligned}
G(1_{B'}, g) \circ G(f, 1_C) \circ \eta_{(B,C)} &= \overbrace{G(1_{B'}, g) \circ \eta_{(B',C)}}^{\text{replace via equation (2)}} \circ F(f, 1_C) \\
&= \eta_{(B',C')} \circ F(1_{B'}, g) \circ F(f, 1_C) \\
&= \eta_{(B',C')} \circ F(1_{B'} \circ f, g \circ 1_C) \\
&= \eta_{(B',C')} \circ F(f, g).
\end{aligned}$$

where in the second step we applied equation (2.2), and in the third step we composed the morphisms. Also note that we can simplify the left-hand side since

$$G(1_{B'}, g) \circ G(f, 1_C) = G(1_{B'} \circ f, g \circ 1_C) = G(f, g).$$

Therefore, we have that

$$G(f, g) \circ \eta_{(B,C)} = \eta_{(B',C')} \circ F(f, g)$$

which implies that *eta* itself is a natural transformation. Specifically, it implies the following diagram.

$$\begin{tikzcd}
(B, C) \arrow[d, "(f,g)"] \\
(B', C')
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
F(B, C) \arrow[r, "\eta_{(B,C)}"] \arrow[d, "F(f,g)"'] & G(B, C) \arrow[d, "G(f,g)"] \\
F(B', C') \arrow[r, "\eta_{(B',C')}"] & G(B', C')
\end{tikzcd}$$

∎

Note: A way to succinctly prove the reverse implication of the previous proof is as follows. Since we know the diagrams on the left are commutative, just "stack" them on top of each other to achieve the diagram in the upper right corner, and then "squish" this diagram down to obtain the third diagram in the bottom right.

$$\begin{tikzcd}
F(B,C) \arrow[r, "\eta_{(B,C)}"] \arrow[d, "F(f,1_C)"'] & G(B,C) \arrow[d, "G(f,1_C)"] \\
F(B',C) \arrow[r, "\eta_{(B',C)}"] & G(B',C)
\end{tikzcd}$$

$$\begin{tikzcd}
F(B',C) \arrow[r, "\eta_{(B',C)}"] \arrow[d, "F(1_{B'},g)"'] & G(B',C) \arrow[d, "G(1_{B'},g)"] \\
F(B',C') \arrow[r, "\eta_{(B',C')}"] & G(B',C')
\end{tikzcd}$$

$$\begin{tikzcd}
F(B,C) \arrow[r, "\eta_{(B,C)}"] \arrow[d, "F(f,1_C)"'] & G(B,C) \arrow[d, "G(f,1_C)"] \\
F(B',C) \arrow[r, "\eta_{(B',C)}"] \arrow[d, "F(1_{B'},g)"'] & G(B',C) \arrow[d, "G(1_{B'},g)"] \\
F(B',C') \arrow[r, "\eta_{(B',C')}"] & G(B',C')
\end{tikzcd}$$

$$\begin{tikzcd}
F(B,C) \arrow[r, "\eta_{(B,C)}"] \arrow[d, "F(f,g)"'] & G(B,C) \arrow[d, "G(f,g)"] \\
F(B',C') \arrow[r, "\eta_{(B',C')}"] & G(B',C')
\end{tikzcd}$$

This is essentially what we did in the proof, although this is more crude visualization of what happened, and we were more formal throughout the process.

## Exercises

**1.** Let $\mathcal{C}$ and $\mathcal{D}$ be categories. Prove that $(\mathcal{C} \times \mathcal{D})^{\mathrm{op}} \cong \mathcal{C}^{\mathrm{op}} \times \mathcal{D}^{\mathrm{op}}$.

## 2.3 Functor Categories

In the proof for the last proposition, we used a trick of forming a desired natural transformation by composing two composable natural transformations. Hence, we see that natural transformations can be "composed." We refine this notion as follows.

Let $\mathcal{C}$ and $\mathcal{D}$ be categories and consider three functors $F, G, H : \mathcal{C} \longrightarrow \mathcal{D}$. Suppose further that we have two natural transformations $\sigma, \tau$ as below:

$$F \xrightarrow{\sigma} G \xrightarrow{\tau} H$$

(This might seem like a weird way to write this, but we are trying to hint at something.) Using these two natural transformations, we can define a natural transformation

$$\tau \cdot \sigma : F \longrightarrow H$$

where, for each $C \in \mathcal{C}$, we define

$$(\tau \cdot \sigma)_C = \tau_C \circ \sigma_C : F(C) \longrightarrow H(C).$$

Visually, we can picture what we are doing as follows. For a given morphism $f : A \longrightarrow B$ in $\mathcal{C}$, we define the morphism $(\tau \cdot \sigma)_C$ as

$$\begin{tikzcd}
F(A) \arrow[r, "F(f)"] \arrow[d, "\sigma_A"'] \arrow[dd, bend right=60, "(\tau \cdot \sigma)_A"'] & F(B) \arrow[d, "\sigma_B"] \arrow[dd, bend left=60, "(\tau \cdot \sigma)_B"] \\
G(A) \arrow[r, "G(f)"] \arrow[d, "\tau_A"'] & G(B) \arrow[d, "\tau_B"] \\
H(A) \arrow[r, "H(f)"] & H(B)
\end{tikzcd}$$

Thus, we see that natural transformations can be "composed," and we can thus ask: If we view functors as objects, and view natural transformations as morphisms, do we get a category? The answer is yes.

**Definition 2.3.1.** Let $\mathcal{C}$ and $\mathcal{D}$ be small categories and consider set of all functors $F : \mathcal{C} \longrightarrow \mathcal{D}$. Then the **functor category**, denoted as $\mathcal{D}^{\mathcal{C}}$ or $\text{Fun}(\mathcal{C}, \mathcal{D})$, is the category where
**Objects.** Functors $F : \mathcal{C} \longrightarrow \mathcal{D}$
**Morphisms.** Natural transformations $\eta : F \longrightarrow G$

Functor categories are extremely useful, as we shall see that they're the categorical version of representations.

When we think of representations, we usually think of a group homomorphism $\rho : G \longrightarrow \text{GL}_n(V)$ for some vector space $V$ over a field $k$. However, suppose we wanted to be a real smartass and say “Well, can’t we regard $\rho$ as actually a functor between two one-object categories whose morphisms are all isomorphism?” The answer is yes!

What this then means is that the category of representations of a group $G$ is actually a functor category. Specifically,

$$\mathrm{Fun}(G, \mathrm{GL}_n(V)) \cong R\text{-}\mathbf{Mod}.$$

Hence in some cases it helps to think of $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ as a category of representations of $\mathcal{C}$. This makes sense, since that is really what a functor is. A functor preserves composition; and if we stop thinking like the set theorists, we can realize that composition controls a great deal of structure in a category $\mathcal{C}$. Hence a functor $F : \mathcal{C} \longrightarrow \mathcal{D}$ “represents"" that structure in a category $\mathcal{D}$.

---

**Example 2.3.2.** Let $\mathbf{1}$ be the one element category with a single identity arrow. Then for any category $\mathcal{C}$, the functor category $\mathcal{C}^{\mathbf{1}}$ is isomorphic to $\mathcal{C}$. This is because each functor $F : \mathbf{1} \longrightarrow \mathcal{C}$ simply associates the element $1 \in \mathbf{1}$ to an element $C \in \mathcal{C}$, and the identity $1_1 : 1 \longrightarrow 1$ to the identity morphism $1_C$ in $\mathcal{C}$.

---

---

**Example 2.3.3.** Let $\mathbf{2}$ be the category consisting of two elements, containing the two identities and one nontrivial morphism between the objects.

$$\begin{tikzcd} 1 \arrow[loop above, "\mathrm{id}_1"] \arrow[r, "f"] & 2 \arrow[loop above, "\mathrm{id}_2"] \end{tikzcd}$$

*The category* $\mathbf{2}$.

Now consider the functor category $\mathcal{C}^{\mathbf{2}}$ where $\mathcal{C}$ is any category. Each functor $F : \mathbf{2} \longrightarrow \mathcal{C}$ maps the pair of objects to objects $F(1)$ and $F(2)$ in $\mathcal{C}$. However, since functors preserve morphisms, we see that

$$f : 1 \longrightarrow 2 \implies F(f) : F(1) \longrightarrow F(2).$$

This is what each $F \in \mathcal{C}^{\mathbf{2}}$ does. Hence, every morphism $g \in \mathrm{Hom}(\mathcal{C})$ corresponds to an element in $\mathcal{C}^{\mathbf{2}}$. Hence, we call $\mathcal{C}^{\mathbf{2}}$ the category of arrows of $\mathcal{C}$.

*Proof.* Let $g : C \longrightarrow C'$ be any morphism between objects $C, C'$ in $\mathcal{C}$. Construct the element $G \in \mathcal{C}^{\mathbf{2}}$ as follows: $G(1) = C$, $G(2) = C'$ and $G(f) : G(1) \longrightarrow G(2) = g$. Hence, $\mathrm{Hom}(\mathcal{C})$ and $\mathcal{C}^{\mathbf{2}}$ are isomorphic. Moreover, $\mathrm{Hom}(\mathcal{C})$ determines the members of $\mathcal{C}^{\mathbf{2}}$.

A crude way to visualize this proof is imaging $1 \longrightarrow 2$ is a "stick" with 1 and 2 on either end, and so the action of any functor is simply taking the stick and applying it to anywhere on the direct graph generated by the category $\mathcal{C}$. Hence, this is why we say $\mathrm{Hom}(\mathcal{C})$ determines the functor category $\mathcal{C}^{\mathbf{2}}$. ∎

---

---

**Example 2.3.4.** Let $X$ be a set. Hence, it is a discrete category, which if recall, it's objects are elements of $X$ and the morphisms are just identity morphisms.

Now consider $\{0,1\}^X$, the category of functors $F : X \longrightarrow \{0,1\}$. Then every functor assigns each element of $x \in X$ to either 0 or 1, and assigns the morphism $1_x : x \longrightarrow x$ to either $1_0 : 0 \longrightarrow 0$ or $1_1 : 1 \longrightarrow 1$.

One way to view this is to consider $\mathcal{P}(X)$, and for each $S \in \mathcal{P}$, assign $x$ to 1 if $x \in S$ or $x$ to 0 if $x \notin S$. All of these mappings may be described by elements of $\mathcal{P}$, but we can also realize that each of these mappings correspond to the functors in $\{0,1\}^X$. Hence, we see that $\{0,1\}^X$ is isomorphic to $\mathcal{P}(X)$.

---

---

**Example 2.3.5.** Recall from Example **??** that, given a group $G$ and a ring $R$ (with identity), we can create a *group ring* $R[G]$ with identity, in a functorial way, establishing a functor

$$R[-] : \mathbf{Grp} \longrightarrow \mathbf{Ring}.$$

However, we then noticed that the above functor establishes a process where we send rings $R$ to functors $R[-] : \mathbf{Grp} \longrightarrow \mathbf{Ring}$. It turns out that this process is itself a functor, and we now have the appropriate language to describe it:

$$F : \mathbf{Ring} \longrightarrow \mathbf{Ring}^{\mathbf{Grp}}$$

Specifically, let $\psi : R \longrightarrow S$ be a ring homomorphism. Now observe that $\psi$ induces another ring homomorphism

$$\psi^*_G : R[G] \longrightarrow S[G] \qquad \sum_{g \in G} a_g g \mapsto \sum_{g \in G} \varphi(a_g) g.$$

As a result, we see that such a ring homomorphism induces a natural transformation. To show this, let $\varphi : G \longrightarrow H$ be a group homomorphism. Then observe that we get the diagram in the middle.

$$\begin{tikzcd} G \arrow[d, "\varphi"] \\ H \end{tikzcd} \qquad \begin{tikzcd} R[G] \arrow[r, "\psi^*_G"] \arrow[d, "R(\varphi)"'] & S[G] \arrow[d, "S(\varphi)"] \\ R[H] \arrow[r, "\psi^*_H"'] & S[H] \end{tikzcd} \qquad \begin{tikzcd} \sum_{g \in G} a_g g \arrow[r, mapsto] \arrow[d, mapsto] & \sum_{g \in G} \psi(a_g) g \arrow[d, mapsto] \\ \sum_{g \in G} a_g \varphi(g) \arrow[r, mapsto] & \sum_{g \in G} \psi(a_g)\varphi(g) \end{tikzcd}$$

However, we can follow the elements as in the diagram on the right, which shows us that the diagram commutes. Hence we see that $\psi^*$ is a natural transformation between functors

$R[-] \longrightarrow S[-]$. Overall, this establishes that we do in fact have a functor

$$F : \mathbf{Ring} \longrightarrow \mathbf{Ring}^{\mathbf{Grp}}$$

which we wouldn't be able to describe without otherwise introducing the notion of a functor category.

---

**Example 2.3.6.** Let $M$ be a monoid category (one object) and consider the functor category $\mathbf{Set}^M$. The objects of $\mathbf{Set}^M$ are functors $F : M \longrightarrow \mathbf{Set}$, each of which have the following data:

$$F(f) : F(M) \longrightarrow F(M)$$

where $f : M \longrightarrow M$ is an morphism in $M$. Now if we interpret $\circ$ as the binary relation equipped on $M$, we see that for any $g : M \longrightarrow M$,

$$F(g \circ f) = F(g) \circ F(f)$$

by functorial properties. Hence, each functor $F$ maps $M$ to a set $X$ which induces the operation of $M$ on $X$. Therefore the objects of $\mathbf{Set}^M$ are other monoids $X$ in $\mathbf{Set}$ equipped with the same operation as $M$ and as well as the morphisms between such monoids.

---

## 2.4 Vertical, Horizontal Composition; Interchange Laws

In the previous section, we considered the idea of forming a composition of natural transformations, and we verified that this formed a valid natural transformation. That is, if we have three functors $F, G, H : \mathcal{C} \longrightarrow \mathcal{D}$ between two categories $\mathcal{C}$ and $\mathcal{D}$, and if $\sigma : F \longrightarrow G$ and $\tau : G \longrightarrow H$ are natural transformations, then we can form the natural transformation

$$(\tau \circ \sigma) : F \longrightarrow H.$$

We call such a type of composition as vertical compositions of natural transformations, since the idea can be captured in the following diagram.

$$\begin{tikzcd} \mathcal{C} \arrow[r, bend left=60, ""{name=U}] \arrow[r, ""{name=M}] \arrow[r, bend right=60, ""{name=L}] & \mathcal{D} \arrow[from=U, to=M, Rightarrow, "\sigma"] \arrow[from=M, to=L, Rightarrow, "\tau"] \end{tikzcd}$$

We can also perform a different, but similar type of composition between natural transformations. Suppose $F, G : \mathcal{B} \longrightarrow \mathcal{C}$ and $F', G' : \mathcal{C} \longrightarrow \mathcal{D}$ are functors between categories $\mathcal{B}, \mathcal{C}$, and $\mathcal{D}$. Furthermore, suppose we have natural transformations $\eta : F \longrightarrow G$ and $\eta' : F' \longrightarrow G'$. Then we have diagram such as the following.

$$\begin{tikzcd} \mathcal{B} \arrow[r, bend left=60, "F", ""{name=A, below}] \arrow[r, bend right=60, "G"', ""{name=B}] & \mathcal{C} \arrow[r, bend left=60, "F'", ""{name=C, below}] \arrow[r, bend right=60, "G'"', ""{name=D}] & \mathcal{D} \arrow[from=A, to=B, Rightarrow, "\eta"] \arrow[from=C, to=D, Rightarrow, "\eta'"] \end{tikzcd}$$

Now let $B$ be an object of $\mathcal{B}$. There are two ways we can transfer this object to an object of $\mathcal{C}$; namely, via mappings of $F$ and $G$. Thus $F(B)$ and $G(B)$ are two objects of $\mathcal{C}$. Since $\eta : F \longrightarrow G$ is a natural transformation between these objects, we see that there's a way of mapping between these two elements in $\mathcal{C}$:

$$\eta(B) : F(B) \longrightarrow G(B).$$

Hence, we have two objects in $\mathcal{C}$ and a morphism in between them. Hence, we know that the natural transformation $\eta' : F' \longrightarrow G'$ implies the following diagram commutes.

$$\begin{tikzcd} F(B) \arrow[d, "\eta(B)"] \\ G(B) \end{tikzcd} \qquad \begin{tikzcd} F' \circ F(B) \arrow[r, "\eta' F(B)"] \arrow[d, "F' \circ \eta(B)"'] & G' \circ F(B) \arrow[d, "G' \circ \eta(B)"] \\ F' \circ G(B) \arrow[r, "\eta' G(B)"] & G' \circ G(B) \end{tikzcd}$$

Note that in the last diagram, all of the objects and morphisms between them exist in $\mathcal{D}$. The easiest way to see why this diagram commutes is to go back directly to the definition of a natural transformation; namely, the pair of objects along with their morphism on the left imply the commutativity of the diagram on the right.

This can be done in general for categories $\mathcal{B}, \mathcal{C}$, and $\mathcal{D}$ which have functors $F, G : \mathcal{B} \longrightarrow \mathcal{C}$ and $F', G' : \mathcal{C} \longrightarrow \mathcal{D}$ associated with natural transformations $\eta : F \longrightarrow G$ and $\eta' : F' \longrightarrow G'$. Furthermore, it holds for all $B \in \mathcal{B}$.

Note further that this diagram is similar to a diagram which represents a natural transformation; but between which functors? If we look closely, we see that it is between $F \circ F'$ and $G \circ G'$.

This leads us to make the following formulaic definition: For natural transformations $\eta : F \longrightarrow G$ and $\eta' : F' \longrightarrow G'$ such that $F, G : \mathcal{B} \longrightarrow \mathcal{C}$ and $F', G' : \mathcal{C} \longrightarrow \mathcal{D}$, then for $B \in \mathcal{B}$ we define their "horizontal" composition as the diagonal of the above diagram; that is,

$$(\eta \circ \eta')B = G'(\eta(B)) \circ \eta' F(B) = \eta'(G(B)) \circ F'(\eta(B)).$$

The above diagram doesn't quite show that $\eta \circ \eta' : F' \circ F \longrightarrow G \circ G'$ is a natural transformation. In order to do this, we need to start from two objects in $\mathcal{B}$ and consider a morphism between them.

**Proposition 2.4.1.** The function $\eta \circ \eta' : F \circ F' \longrightarrow G \circ G'$ is a natural transformation between the functors $F' \circ F, G' \circ G : \mathcal{B} \longrightarrow \mathcal{D}$.

*Proof.* To show this, we consider a morphism $f : B \longrightarrow B'$ between two objects $B$ and $B'$ in $\mathcal{B}$. We then claim that the following diagram is commutative:

$$\begin{tikzcd}
B \arrow[d, "f"] & F' \circ F(B) \arrow[r, "F' \circ \eta(B)"] \arrow[d, "F' \circ F(f)"'] & F' \circ G(B) \arrow[r, "\eta' \circ G(B)"] \arrow[d, "F' \circ G(f)"] & G' \circ G(B) \arrow[d, "G' \circ G(f)"] \\
B' & F' \circ F(B') \arrow[r, "F' \circ \eta(B')"'] & F' \circ G(B') \arrow[r, "\eta' \circ \eta(B')"'] & G' \circ G(B')
\end{tikzcd}$$

First, observe that the left square is commutative due to the fact that $\eta$ is a natural transformation from $F$ to $G$. Therefore, it produces a commutative square diagram, and we obtain the above left square diagram by applying $F'$ to the commutative diagram produced by $\eta : F \longrightarrow G$.

The right square in the diagram is obtained by the fact that $\eta'$ is a natural transformation between functors $F'$ and $G'$. Hence the diagram is commutative, and it acts on the objects $G(B)$ and in $\mathcal{C}$. Therefore, we see that $\eta \circ \eta'$ is a natural transformation. ■

Thus we see that we have "horizontal" and "vertical" notions of composing natural transformations. Let us denote "horizontal" transformations as $\circ$ and "vertical" transformations as $\cdot$ between natural transformations.

It is also notationally convenient to denote functor and natural transformation compositions as

$$F' \circ \tau : F' \circ F \longrightarrow F' \circ T \quad \eta' \circ G : F' \circ G \longrightarrow G' \circ G$$

which are two additional natural transformations. (Remember we showed that the left square in the commutative diagram of the previous proof commuted by observing that it was obtained by the commutative diagram produced by the natural transformation $\eta$ and composing it with $F'$? What we really showed is that $F' \circ \eta$ is a natural transformation, since this natural transformation described that square. Similarly, $\eta' \circ G$ is the natural transformation which represents the right square of the commutative diagram in the previous proof.)

With the above notation, we can then write that

$$\eta' \circ \eta = (G' \circ \eta) \cdot (\eta' \circ F) = (\eta' \circ G) \cdot (F' \circ \eta).$$

This idea of ours can be extended to a more general situation. Suppose we have instead three categories $\mathcal{B}, \mathcal{C}$, and $\mathcal{D}$ and where $F, G, H : \mathcal{B} \longrightarrow \mathcal{C}$ and $F, G, H : \mathcal{C} \longrightarrow \mathcal{D}$ are functors associated with natural transformations $\eta : F \longrightarrow G, \sigma : G \longrightarrow H$, and $\eta' : F' \longrightarrow G', \sigma' : G' \longrightarrow H'$. The following diagram may be more helpful than words:

$$\begin{tikzcd}[column sep=large]
\mathcal{B} \arrow[r, shift left=3ex, "F"] \arrow[r] \arrow[r, shift right=3ex, "H"'] \arrow[r, phantom, shift left=1.5ex, "\Downarrow \eta"] \arrow[r, phantom, shift right=1.5ex, "\Downarrow \sigma"] & \mathcal{C} \arrow[r, shift left=3ex, "F'"] \arrow[r] \arrow[r, shift right=3ex, "H'"'] \arrow[r, phantom, shift left=1.5ex, "\Downarrow \eta'"] \arrow[r, phantom, shift right=1.5ex, "\Downarrow \sigma'"] & \mathcal{D}
\end{tikzcd}$$

Note we've omitted the label of $G$ and $G'$ on the middle horizontal arrows since they don't exactly fit in there when we include the labels for the natural transformations.

Now suppose we have an object $B$ in $\mathcal{B}$. Then we can create three objects $F(B), G(B)$ and $H(B)$ in $\mathcal{C}$, and we may interchange between these objects via the given natural transformations. Specifically, $\eta(B) : F(B) \longrightarrow G(B)$ and $\sigma(B) : G(B) \longrightarrow H(B)$. However, we also know that $\eta', \sigma'$ are natural transformations between $\mathcal{C}$ and $\mathcal{D}$, and hence imply the following commutative diagram.

$$\begin{tikzcd}
F(B) \arrow[d, "\eta(B)"] & F' \circ F(B) \arrow[r, "\eta' F(B)"] \arrow[d, "F' \circ \eta(B)"'] & G' \circ F(B) \arrow[r, "\sigma' F(B)"] \arrow[d, "G' \circ \eta(B)"] & H' \circ F(B) \arrow[d, "H' \circ \eta(B)"] \\
G(B) \arrow[d, "\sigma(B)"] & F' \circ G(B) \arrow[r, "\eta' G(B)"] \arrow[d, "F' \circ \sigma(B)"'] & G' \circ G(B) \arrow[r, "\sigma' G(B)"] \arrow[d, "G' \circ \sigma(B)"] & H' \circ G(B) \arrow[d, "H' \circ \sigma(B)"] \\
H(B) & F' \circ H(B) \arrow[r, "\eta' H(B)"'] & G' \circ H(B) \arrow[r, "\sigma' H(B)"'] & H' \circ H(B)
\end{tikzcd}$$

Suppose we start at the upper left corner and want to achieve the value at the bottom right. There are two ways we can do this: We can travel within the interior of the diagram, or we can travel on the outside of the diagram.

In traveling on the interior of the diagram, note that the composition of the arrows of the upper left square is $\eta' \circ \eta$. In addition, composition of the arrows of the bottom right square is $\sigma' \circ \sigma$.

In traveling on the exterior of the diagram note that the composition of the top row is $\eta' \cdot \sigma'$ and composition of the right most vertical arrows is $\eta \cdot \sigma$. Since both paths achieve the same value, we see that

$$(\eta' \cdot \sigma') \circ (\eta \cdot \sigma) = (\eta' \circ \eta) \cdot (\sigma' \circ \sigma)$$

which is known as the **Interchange Law**.

This leads us to make the following definition.

**Definition 2.4.2.** We define a **double category** to be a set of arrows which obey two different forms of composition, generally denoted as $\circ$ and $\cdot$, which together satisfy the interchange law.

Furthermore, a **2-category** is a double category in which $\cdot$ and $\circ$ have the same exact identity arrows.

## 2.5 Slice and Comma Categories.

In this section we introduce comma categories, which serve as a very useful categorical construction. The reason why it is so useful is because the notion of a comma category has the potential to simplify an otherwise complicated discussion. As they can be constructed in any category, and because they contain a large amount of useful data, they are frequently used as an intermediate step in more complex categorical constructions. Thus, while the concept is "simple," they nevertheless appear in all kinds of complicated discussions in category theory.

**Definition 2.5.1.** Let $\mathcal{C}$ be a category and suppose $A$ is an object of $\mathcal{C}$. We define the **slice category (with $A$ over $\mathcal{C}$)**, denoted $(A \downarrow \mathcal{C})$, as the category

**Objects.** All pairs $(C, f : A \longrightarrow C)$ for all $C \in \mathcal{C}$ and morphims $f : A \longrightarrow C$. In other words, the objects are all morphisms in $\mathcal{C}$ which *originate* at $A$.

**Morphisms.** For two objects $(C, f : A \longrightarrow C)$ and $(C', f' : A \longrightarrow C')$, we define

$$h : (C, f) \longrightarrow (C', f')$$

as a morphism between the objects, where $h : C \longrightarrow C'$ is a morphism in our category such that $f' = h \circ f$. Alternatively we can describe the homset more directly:

$$\text{Hom}_{(A\downarrow\mathcal{C})}\Big((f, C), (f', C')\Big) = \{h : C \longrightarrow C' \in \mathcal{C} \mid f' = h \circ f\}.$$

At this point you may be a bit overloaded with notation if this is the first time you've seen this before. You need to figure out how this is a category (what's the identity? composition?) and ultimately why you should care about this category. To aid your understanding, a picture might help.

We can represent the objects and morphisms of the category $(A \downarrow \mathcal{C})$ in a visual manner.

Objects $(C, f)$:
$$\begin{tikzcd} A \arrow[d, "f"] \\ C \end{tikzcd}$$

Morphisms $h : (C, f) \longrightarrow (C', f')$ are given by $h : C \longrightarrow C'$ such that
$$\begin{tikzcd} & A \arrow[dl, "f"'] \arrow[dr, "f'"] & \\ C \arrow[rr, "h"'] & & C' \end{tikzcd}$$

Now, how does composition work? Composition of two composable morphisms $h : (f, C) \longrightarrow (f', C')$ and $h' : (f', C') \longrightarrow (f'', C'')$ is given by $h' \circ h : (f, C) \longrightarrow (f'', C'')$ since clearly

$$f'' = h' \circ f' \quad \text{and} \quad f' = h \circ f \implies f'' = h' \circ (h \circ f) = (h' \circ h) \circ f.$$

We can visually justify composition as well. If we have two commutative diagrams as on the left, we can just squish them together to get the final commutative diagram on the right.

$$\begin{tikzcd} & A \arrow[dl, "f"'] \arrow[dr, "f'"] & \\ C \arrow[rr, "h"'] & & C' \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} & A \arrow[dl, "f'"'] \arrow[dr, "f''"] & \\ C' \arrow[rr, "h'"'] & & C'' \end{tikzcd} \quad \text{implies} \quad \begin{tikzcd} & A \arrow[dl, "f"'] \arrow[dr, "f''"] & \\ C \arrow[rr, "h' \circ h"'] & & C'' \end{tikzcd}$$

Hence, we see that $h' \circ h : (f, C) \longrightarrow (f'', C'')$ is defined whenever $h'$ and $h$ are composable as morphisms of $\mathcal{C}$.

One use of comma categories is to capture and generalize the notion of a pointed category. Such pointed categories include the category of pointed sets $\mathbf{Set}^*$ or the category of pointed topological spaces $\mathbf{Top}^*$, etc.

We've seen, in particular on the discussion of functors, the necessity for pointed categories. For example, we cannot discuss "the" fundamental group $\pi_1(X)$ of a topological space $X$ (unless $X$ is path connected, but still only up to isomorphism). To discuss a fundamental group in a topological space $X$, one needs to select a base point $x_0$. As we saw in Example 1.7, $\pi_1$ is not a functor $\mathbf{Top} \longrightarrow \mathbf{Grp}$, but is rather a functor

$$\pi_1 : \mathbf{Top}^* \longrightarrow \mathbf{Grp}$$

where $\mathbf{Top}^*$, which consists of pairs $(X, x_0)$ with $x_0 \in X$, is the category of pointed topological spaces.

Similarly, it makes no sense to talk about "the" tangent plane of a smooth manifold. Such an association requires the selection of a point $p \in X$ to calculate $T_p(M)$. So, as we saw in Example **??**, this process is not a functor from **DMan** to **Vect**, but is rather a functor

$$T : \mathbf{DMan}^* \longrightarrow \mathbf{Vect}$$

where $\mathbf{DMan}^*$, which consists of pairs $(M, p)$ with $p \in M$, is the category of pointed smooth manifolds. This now motivates the next two examples.

---

**Example 2.5.2.** Consider the category $\mathbf{Top}^*$ where

**Objects.** The objects are pairs $(X, x_0)$ with $X$ a topological space and $x_0 \in X$.

**Morphisms.** A morphism $f : (X, x_0) \longrightarrow (Y, y_0)$ is any continuous function $f : X \longrightarrow Y$ such that $y_0 = f(x_0)$.

Recall that the one point set $\{\bullet\}$ is trivially a topological space. Then we can form the category $(\{\bullet\} \downarrow \mathbf{Top})$. The claim now is that

$$(\{\bullet\} \downarrow \mathbf{Top}) \cong \mathbf{Top}^*.$$

Why? Well, an object of $(\{\bullet\} \downarrow \mathbf{Top})$ is simply a pair $(X, f : \{\bullet\} \longrightarrow X)$. Observe that

$$f(\bullet) = x_0 \in X,$$

for some $x_0 \in X$. So, the pair $(X, f : \{\bullet\} \longrightarrow X)$ is logically equivalent to a pair $(X, x_0)$ with

$x_0 \in X$. That is, a continuous function from the one point set into a topological space $X$ is equivalent to simply selecting a single point $x_0 \in X$. Hence, on objects it is clear why we have an isomorphism.

Now, a morphism in this comma category will be of the form $p : (X, f_1 : \{\bullet\} \longrightarrow X) \longrightarrow (Y, f_1 : \{\bullet\} \longrightarrow Y)$. Specifically, it is a continuous function $p : X \longrightarrow Y$ such that the diagram below commutes.

$$\begin{tikzcd} & \{\bullet\} \arrow[dl, "f_1"'] \arrow[dr, "f_2"] & \\ X \arrow[rr, "p"'] & & Y \end{tikzcd}$$

In other words, if $f_1(\bullet) = x_0$ and $f_2(\bullet) = y_0$, it is a continuous function $p : X \longrightarrow Y$ such that $f(x_0) = y_0$. This is exactly a morphism in $\textbf{Top}^*$! We clearly have a bijection as claimed.

---

The above example generalizes to many pointed categories, some of which are

- $\textbf{DMan}^* \cong (\bullet \downarrow \textbf{DMan})$
- $\textbf{Set}^* \cong (\bullet \downarrow \textbf{Set})$
- $\textbf{Grp}^* \cong (\bullet \downarrow \textbf{Grp})$

We now briefly comment for any slice category $(A \downarrow \mathcal{C})$ built from a category $\mathcal{C}$, we can construct a "projection" functor

$$P : (A \downarrow \mathcal{C}) \longrightarrow \mathcal{C}$$

where on objects $P(C, f : A \longrightarrow C) = C$ and on morphisms $P(h : (C, f) \longrightarrow (C', f')) = h : C \longrightarrow C'$. Clearly, this functor is faithful, but it is generally not full. Such a projection functor is used in technical constructions involving slice categories as it has nice properties; we will make use of it later when we discuss limits.

Next,we introduce how we can also describe the category of an objects *under* another category.

**Definition 2.5.3.** Let $\mathcal{C}$ be a category, and $B$ an object of $\mathcal{C}$. Then we define the **category** $B$ **under** $\mathcal{C}$, denoted as $(\mathcal{C} \downarrow B)$ as follows.

**Objects.** All pairs $(C, f)$ where $f : C \longrightarrow B$ is a morphism in $\mathcal{C}$. That is, the objects are morphisms *ending* at $B$.

**Morphisms.** For two objects $(C, f : C \longrightarrow B)$ and $(C', f' : C' \longrightarrow B)$, we define

$$h : (C, f) \longrightarrow (C', f')$$

to be a morphism between the objects to correspond to a morphism $h : C \longrightarrow C'$ in $\mathcal{C}$ such that $f = f' \circ h$.

Composition of functions $h : (f, C) \longrightarrow (f', C')$ and $h' : (f', C') \longrightarrow (f'', C'')$ exists whenever $h' \circ h$ is defined as morphisms in $\mathcal{C}$. Again, we can represent the elements of the category in a visual manner

$$\text{Objects } (C, f): \quad \begin{tikzcd} C \arrow[d, "f"] \\ B \end{tikzcd} \qquad \text{Morphisms } h : (f, C) \longrightarrow (f', C') \quad \begin{tikzcd} C \arrow[rr, "h"] \arrow[dr, "f"'] & & C' \arrow[dl, "f'"] \\ & B & \end{tikzcd}$$

The following is a nice example that isn't traditionally seen as an example of a functor.

---

**Example 2.5.4.** Let $(G, \cdot)$ and $(H, \cdot)$ be two groups, and consider a group homomorphism $\varphi : (G, \cdot) \longrightarrow (H, \cdot)$. Abstractly, this is an element of the comma category $(\mathbf{Grp} \downarrow H)$.

Now for for every group homomorphism, we may calculate the kernal of $\text{Ker}(\varphi) = \{g \in G \mid \varphi(g) = 0\}$. This is always a subgroup of $G$. What is interesting is that, from the perspective of slice categories, this process is functorial:

$$\text{Ker}(-) : (\mathbf{Grp} \downarrow H) \longrightarrow \mathbf{Grp}.$$

To see this, we have to understand what happens on the morphisms. So, suppose we have two objects $(G, \varphi : G \longrightarrow H)$ and $(K, \psi : K \longrightarrow H)$ of $(\mathbf{Grp} \downarrow H)$ and a morphism $h : G \longrightarrow K$ between the objects.

$$\begin{tikzcd} G \arrow[rr, "h"] \arrow[dr, "\varphi"'] & & K \arrow[dl, "\psi"] \\ & H & \end{tikzcd}$$

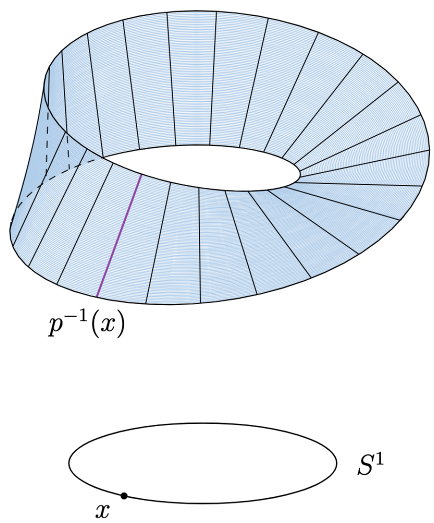
Then we can define $\text{Ker}(h) : \text{Ker}(\varphi) \longrightarrow \text{Ker}(\psi)$, the image of $h$ under the functor, to be the restriction $h|_{\ker(\varphi)} : \text{Ker}(\varphi) \longrightarrow \text{Ker}(\psi)$. This is a bonafied group homomorphism: by the commutativity of the above triangle, if $g \in G$ then $\varphi(g) = \psi(h(g))$. Hence, if $\varphi(g) = 0$, i.e., $g \in \text{Ker}(\varphi)$, then $\psi(h(g)) = 0$, i.e., $h(g) \in \text{Ker}(\psi)$. So we see that our proposed function makes sense.

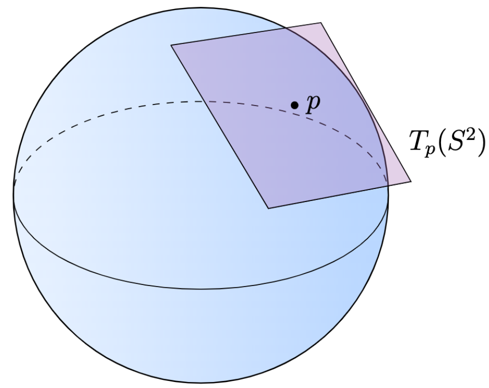
What this means is that the commutativity of the above triangle forces a natural relationship between the kernels of $\varphi$ and $\psi$; not only as a function of sets, but as a group homomorphism. Therefore, the kernel of a group homomorphism is actually a functor from a slice category.

---

---

**Example 2.5.5.** In geometry and topology, one often meets the need to define a $(-)$-bundle. By $(-)$ we mean vector, group, etc. That is, we often want topological spaces to parameterize a family of vector spaces or groups in a coherent way.




For example, on the above left we can map the Möbius strip onto $S_1$ in such a way that the inverse image of each $x \in S_1$ is homeomorphic to the interval $[0, 1]$. Hence, each point of $x \in S_1$ carries the information of a topological space, specifically one of $[0, 1]$.

On the right, we can recall that $S^2$ is a differentiable manifold, and so each point $p$ has a tangent plane $T_p(S^2)$, which is a vector space. Hence every point on $S^2$, or more generally for any differentiable manifold, carries the information of a vector space.

In general, for a topological space $X$, we define a **bundle** over $X$ to be a continuous map $p : E \longrightarrow X$ with $E$ being some topological space of interest. If $p : E \longrightarrow X$ an $p' : E' \longrightarrow X$ are two bundles, a **morphism of bundles** $q : p \longrightarrow p'$ is given by a continuous map $q : E \longrightarrow E'$ such that

$$p = p' \circ q.$$

Hence we see that a bundle over a topological space $X$ is an element of the comma category $\mathbf{Top}/X$, and a morphism of bundles is a morphism in the comma category. We therefore see that $\mathbf{Top}/X$ can be interpreted as the **category of bundles of** $X$.

One particular case of interest concerns **vector bundles.** Let $E, X$ be topological spaces. Recall that a vector bundle consists of a continuous map $\pi : E \longrightarrow X$ such that

1. $\pi^{-1}(x)$ is a finite-dimensional vector space over some field $k$

2. For each $p \in X$, there is an open neighborhood $U_\alpha$ and a homeomorphism

$$\varphi_\alpha : U_\alpha \times \mathbb{R}^n \xrightarrow{\sim} \pi^{-1}(U_\alpha)$$

   with $n$ some natural number. We also require that $\pi \circ \varphi_\alpha = 1_{U_\alpha}$.

As we might expect, a **morphism of vector bundles** between $\pi_1 : E \longrightarrow X$ and $\pi_2 : E' \longrightarrow X$ is given by a continuous map $q : E \longrightarrow E'$ such that for each $x \in X$, $q\big|_{\pi_1^{-1}(x)} : \pi_1^{-1}(x) \longrightarrow \pi_2^{-1}(x)$ is linear map between vector spaces.

To realize this in real mathematics, we can take the classic example of the **tangent bundle** on a smooth manifold $M$ (if you've seen this before, hopefully it is now clear why the word "bundle" is here). In differential geometry this is defined as the set

$$TM = \{(p, v) \mid p \in M \text{ and } v \in T_p(M)\}$$

where we recall that $T_p(M)$ is the tangent (vector) space at a point $p \in M$. Since $M$ is a smooth manifold there is a differentiable structure $(U_\alpha, \boldsymbol{x}_\alpha : U_\alpha \longrightarrow M)$ which allow us to define a map

$$\boldsymbol{y}_\alpha : U_\alpha \times \mathbb{R}^n \longrightarrow TM$$
$$((x_1, \dots, x_n), (u_1, \dots, u_n)) \mapsto \left(\boldsymbol{x}_{\boldsymbol{\alpha}}(x_1, \dots, x_n), \sum_{i=1}^{n} u_i \frac{\partial}{\partial x_i}\right).$$

This actually provides a differentiable structure on $TM$, demonstrating it too is a smooth manifold (see Do Carmo). Hence we see that $TM$ is in fact a topological space. We then see that the mapping $\pi : TM \longrightarrow M$ where

$$\pi(p, v) = p \text{ and } \pi^{-1}(x) = T_x(M).$$

is a continuous mapping. Hence we've satisfied both (1.) and (2.) in the the definition of a vector bundle. The other properties can be easily verified so that this provides a nice example of a vector bundle.

---

We can also formulate categories of objects *under* and *over* functors.

**Definition 2.5.6.** Let $\mathcal{C}$ be a category, $C$ an object of $\mathcal{C}$ and $F : \mathcal{B} \longrightarrow \mathcal{C}$ a functor. Then we define the **category** $C$ **over the functor** $F$, denoted as $(C \downarrow F)$, as follows.

**Objects.** All pairs $(f, B)$ where $B \in \text{Obj}(\mathcal{B})$ such that

$$f : C \longrightarrow F(B)$$

where $f$ is a morphism in $\mathcal{C}$.

**Morphisms.** The morphisms $h : (f, B) \longrightarrow (f', B')$ of $(C \downarrow F)$ are defined whenever there exists a $h : B \longrightarrow B'$ in $\mathcal{B}$ such that $f' = F(h) \circ f$.

Representing this visually, we have that

Objects $(f, B)$:
$$\begin{tikzcd} C \arrow[d, "f"] \\ F(B) \end{tikzcd}$$

Morphisms $h : (f, B) \longrightarrow (f', B')$
$$\begin{tikzcd} & C \arrow[dl, "f"'] \arrow[dr, "f'"] & \\ F(B) \arrow[rr, "F(h)"'] & & F(B') \end{tikzcd}$$

Composition of the morphisms in $(C \downarrow F)$ simply requires composition of morphisms in $\mathcal{B}$.

One can easily construct the **category** $C$ **under the functor** $F$, $(F \downarrow C)$, in a completely analogous manner as before. But we'll move onto finally defining the concept of the comma category.

**Definition 2.5.7.** Let $\mathcal{B}, \mathcal{C}, \mathcal{D}$ be categories and let $F : \mathcal{B} \longrightarrow \mathcal{D}$ and $G : \mathcal{C} \longrightarrow \mathcal{D}$ functors. That is,

$$\mathcal{B} \xrightarrow{F} \mathcal{D} \xleftarrow{G} \mathcal{C}.$$

Then we define the **comma category** $(F \downarrow G)$ as follows.

**Objects.** All pairs $(B, C, f)$ where $B, C$ are objects of $\mathcal{B}, \mathcal{C}$, respectively, such that

$$f : F(B) \longrightarrow G(C)$$

where $f$ is a morphism in $\mathcal{D}$.

**Morphisms.** All pairs $(h, k) : (B, C, f) \longrightarrow (B', C', f')$ where $h : B \longrightarrow B'$ and $k : C \longrightarrow C'$ are morphisms in $\mathcal{B}, \mathcal{C}$, respectively, such that

$$f' \circ F(h) = G(k) \circ f.$$

As usual, we can represent this visually via diagrams:

Objects $(B, C, f)$:
$$\begin{tikzcd} F(B) \arrow[d, "f"] \\ G(C) \end{tikzcd}$$

Morphisms $(h, k)$
$$\begin{tikzcd} F(B) \arrow[r, "F(h)"] \arrow[d, "f"'] & F(B') \arrow[d, "f'"] \\ G(C) \arrow[r, "G(k)"'] & G(C') \end{tikzcd}$$

where in the above picture we have that $(h, k) : (B, C, f) \longrightarrow (B', C', f')$. Since functors naturally respect composition of functions, one can easily define composition of morphism $(h, k)$ and $(h', k')$ as $(h \circ h', k \circ k')$ whenever $h \circ h'$ and $k \circ k'$ are defined as morphisms in $\mathcal{B}$ and $\mathcal{C}$, respectively.

## Exercises

**1.** Let $\mathcal{C}$ be a category with initial and terminal objects $I$ and $T$.

*i.* Show that $(\mathcal{C} \downarrow T) \cong \mathcal{C}$.

*i.* Also show that $(I \downarrow \mathcal{C}) \cong \mathcal{C}$.

**2.** Consider again a group homomorphism $\varphi : G \longrightarrow H$, but this time consider the image $\text{Im}(\varphi) = \{\varphi(g) \mid g \in G\}$. Show that this defines a functor

$$\text{Im}(-) : (G \downarrow \textbf{Grp}) \longrightarrow \textbf{Grp}$$

where on morphisms, a morphism

$$h : (H, \varphi : G \longrightarrow H) \longrightarrow (K, \varphi : G \longrightarrow K)$$

is mapped to the restriction $h|_{\text{Im}(\varphi)} : \text{Im}(\varphi) \longrightarrow \text{Im}(\psi)$.

In some sense, this is the "opposite" construction of the kernel functor we introduced. Instead of taking the kernel of a group homomorphism, we can take its image.

**3.** Here we prove that the processes of imposing the **induced topology** and the **coinduced topology** are functorial. Moreover, the correct language to describe this is via slice categories.

*i.* Let $X$ be a set and $(Y, \tau)$ a topological space. Denote $U$ : **Top** $\longrightarrow$ **Set** to be the forgetful functor. Given any function $f : X \longrightarrow U(Y)$, we can use the topology on $Y$ to impose a topology $\tau_X$ on $X$:

$$\tau_X = \{U \subseteq X \mid f(U) \text{ is open in } Y\}.$$

This is called the **induced topology on** $X$. So, we see that (by abuse of notation) the function $f : X \longrightarrow U(Y)$ is now a continuous function $f : (X, \tau_X) \longrightarrow (Y, \tau_Y)$.

Prove that this process forms a functor Ind : $(\mathbf{Top} \downarrow U(Y)) \longrightarrow (\mathbf{Top} \downarrow Y)$.

*ii.* This time, let $(X, \tau)$ be a topological space, $Y$ a set, and consider a function $f : U(X) \longrightarrow Y$. We can similarly impose a topology $\tau_Y$ on $Y$:

$$\tau_Y = \{V \subseteq Y \mid f^{-1}(V) \text{ is open in } X\}.$$

This is called the **coinduced topology on** $Y$. Show that this is also a functorial process.

**4.** *i.* Let $X$, $Y$ be topological spaces with $\varphi : X \longrightarrow Y$ a continuous function. Show that this induces a functor $\varphi_* : (\mathbf{Top} \downarrow X) \longrightarrow (\mathbf{Top} \downarrow Y)$ where on objects $(f : E \longrightarrow X) \mapsto (\varphi \circ f : E \longrightarrow Y)$.

*ii.* Let $\mathcal{C}$ be a category. Show that we generalize $(i)$ to define a functor

$$(\mathcal{C} \downarrow -) : \mathcal{C} \longrightarrow \mathbf{Cat}$$

where $A \mapsto (\mathcal{C} \downarrow A)$.

*ii.* Let $\mathbf{Cat}_*$ be the **pointed category of categories** which we describe as

**Objects.** All pairs $(\mathcal{C}, A)$ with $\mathcal{C}$ a category and $A \in C$
**Morphisms.** Functors $F$ which preserve the objects.

Can we overall describe the construction of a slice category as a functor

$$(- \downarrow -) : \mathbf{Cat}_* \longrightarrow \mathbf{Cat}$$

where $(\mathcal{C}, A) \mapsto (\mathcal{C} \downarrow A)$?

**5.** In this exercise we'll see that slice categories describe intervals for thin categories.

*i.* Regard $\mathbb{R}$ as a thin category, specifically as one with a partial order. For a given $a \in \mathbb{R}$, describe the thin category $(a \downarrow \mathbb{R})$.

*ii.* Suppose $P$ is a partial order (so that $p \leq p'$ and $p' \leq p$ implies $p = p'$). Describe in general the categories $(p \downarrow P)$ and $(P \downarrow p)$.

## 2.6 Graphs, Quivers and Free Categories

In studying category it is often helpful to imagine the objects and morphisms in action as vertices and edges corresponding to a graph. In fact, such a pictorial representation of a category is not even incorrect; one can pass categories and graphs from one to the other. To speak of this, we first review some terminology.

**Definition 2.6.1.** A (small) **graph** $G$ is a set of vertices $V(G)$ and a set edges $E(G)$ such that there exists an assignment function

$$\partial : E(G) \longrightarrow V(G) \times V(G)$$

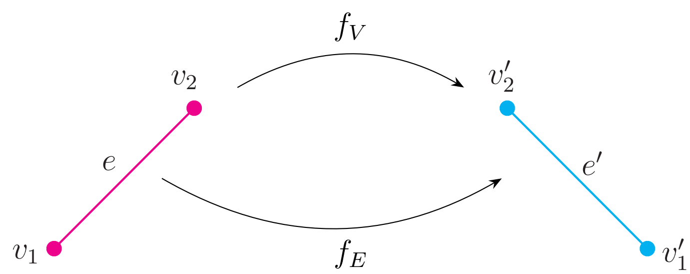
which assigns every edge to the ordered pair containing its endpoints.

On the other hand, a **directed graph** is a graph $G$ where $E(G)$ is now a set of 2-tuples $(v_1, v_2)$. This allows each edge of $E(G)$ to have a specified direction. In this case, the assignment function has the form $\partial : E(G) \longrightarrow V(G)$.

Now, how do we formulate a morphism between two graphs?

**Definition 2.6.2.** A **graph homomorphism** between two graphs $G$ and $H$ is a function $f : G \longrightarrow H$ which induces maps $f_V : V(G) \longrightarrow V(H)$ and $f_E : E(G) \longrightarrow E(H)$ where if $\partial(e) = (v_1, v_2)$, then

$$\partial \circ f_E(e) = (f_V(v_1), f_V(v_2)).$$




In some sense, this behaves almost like a functor. This observation will become important later. Now since we have a consistent way to speak of graphs and their morphisms, we can form the category **Grph** where the objects are small graphs and the morphisms are graph morphisms as described above.

Finally we introduce the concept of a quiver, which we will see is basically the skeleton of a category.

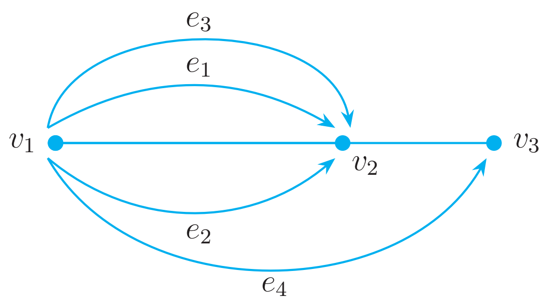
**Definition 2.6.3.** A **quiver** is a directed graph $G$ which allows multiple edges between vertices. Instead of a function $\partial$, a quiver is equipped with **source** and **target** functions

$$s : E(G) \longrightarrow V(G) \qquad t : E(G) \longrightarrow V(G).$$

So a quiver is a 4-tuple $(E(G), V(G), s, t)$. Now as before, a **morphism** $f : Q \longrightarrow Q'$ between quivers $(E(Q), V(Q), s, t)$ and $(E(Q'), V(Q'), s', t')$ is one which preserves edge-vertex relations.

Thus, it is a pair of functions $f_E : E(Q) \longrightarrow E(Q')$ and $f_V : V(Q) \longrightarrow V(Q)'$ such that

$$f_V \circ s = s' \circ f_E \qquad f_V \circ t = t' \circ f_E.$$




Now that we have all of those definitions out of the way, what's really going on here? A quiver can be abstracted as a pair of objects and morphisms.

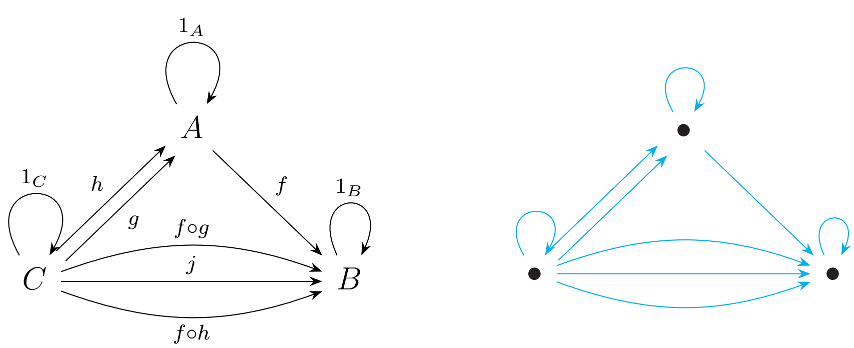
$$E \underset{t}{\overset{s}{\rightrightarrows}} V$$

If we let $C^{\text{op}}$ be the category with two objects, two nontrivial morphisms and two identity morphisms as below

$$1 \underset{g}{\overset{f}{\rightrightarrows}} 0$$

then we see that a **quiver is a functor** $F : \mathcal{C}^{\text{op}} \longrightarrow \mathbf{Set}$. With that said, we can define the **category of quivers Quiv**, which, based on what we just showed, is a functor category with objects $F : \mathcal{C}^{\text{op}} \longrightarrow \mathbf{Set}$. This allows us to interpret quiver homomorphisms as natural transformations.

Now why on earth do we care about these things called quivers? The reason is because the underlying structure of small categories take the form of a quiver. For example, the category on the left below can be turned into a quiver, as on the right, after "forgetting" composition and identity morphisms.




In general, since categories allow multiple arrows between objects, we can construct a forgetful functor which forgets composition and identity arrows.

$$U : \mathbf{Cat} \longrightarrow \mathbf{Quiv}.$$

Note that if $F : \mathcal{C} \longrightarrow \mathcal{C}'$ is a functor then $U(F) : U(\mathcal{C}) \longrightarrow U(\mathcal{C}')$ is in fact a well-behaved morphism between two quivers. Recall that the construction of a graph homomorphism is basically a functor as we've known to so far.

Not only can we forget categories to generate quivers, we can generate categories using the skeletal structure of a quiver. This leads to the concept of a **free category**; the concept is no different than the concept of, say, a free group generated by a set $X$.

**Definition 2.6.4.** Let $Q$ be a quiver with vertex set $V$ and edge set $E$. We define the **free category generated by** $Q$ as the category with

**Objects.** The set $V$

**Morphisms.** The **paths** of the quiver.

Precisely, a **path** is any sequence of edges and vertices

$$v_0 \xrightarrow{e_0} v_1 \xrightarrow{e_1} \cdots \xrightarrow{e_{n-1}} v_n$$

with composition of paths defined in the intuitive way:

$$(v_0 \xrightarrow{e_0} v_1 \xrightarrow{e_1} \cdots \xrightarrow{e_{n-1}} v_n) \circ (v_n \xrightarrow{e'_0} w_0 \xrightarrow{e'_1} w_1 \xrightarrow{e'_2} \cdots \xrightarrow{e'_m} w_m)$$

$$= v_0 \xrightarrow{e_0} v_1 \xrightarrow{e_1} \cdots \xrightarrow{e_{n-1}} v_n \xrightarrow{e'_0} w_0 \xrightarrow{e'_1} w_1 \xrightarrow{e'_2} \cdots \xrightarrow{e'_m} w_m$$

When we generate the free category, we also remember to add identity arrows to each vertex.

Since for each quiver $Q$, we can define a free category $F_C(Q)$ on $Q$, we can realize that this mapping is functorial. That is, we can define a functor

$$F_C : \mathbf{Quiv} \longrightarrow \mathbf{Cat}$$

where it maps on objects and morphisms as

$$\begin{gathered} Q \longmapsto F_C(Q) \\ (f : Q \longrightarrow Q') \longmapsto (F_C(f) : F_C(Q') \longrightarrow F_C(Q)). \end{gathered}$$

That is, quiver homomorphisms can map to functors $F_C(f)$ between the free categories generated by the respective quivers.

Now, what is the relationship between a quiver $Q$ and the quiver $U(F_C(Q))$? There must exist an injection $i : Q \longrightarrow U(F(Q))$ which sends $Q$ to the skeleton of $U(F_C(Q))$. It turns out that this morphism is universal from $Q$ to $U$.

**Theorem 2.6.5.** Let $Q$ be a quiver. Then there is a graph homomorphism $i : Q \longrightarrow U(F_C(Q))$ such that, for any other graph homomorphism $\varphi : Q \longrightarrow U(\mathcal{C})$ with $\mathcal{C}$ a category, there exists a unique functor $F : F_C(Q) \longrightarrow \mathcal{C}$ where $U(F) \circ i = \varphi$. That is,

$$\begin{tikzcd} Q \arrow[r, "i"] \arrow[dr, "\varphi"'] & U(F_C(Q)) \arrow[d, dashed, "U(F)"] & F_C(Q) \arrow[d, dashed, "F"] \\ & U(\mathcal{C}) & \mathcal{C} \end{tikzcd}$$

This is an example of a universal arrow; the dotted lines are the morphisms which are forced to exist by the conditions of the diagram, which is the idea of a universal element.

*Proof.* Denote each morphism or path in $F_C(Q)$ of length $n$

$$v_0 \xrightarrow{e_0} v_1 \xrightarrow{e_1} \cdots \xrightarrow{e_{n-1}} v_n$$

as $(v_0, e_0e_1 \cdots e_{n-1}, v_n) : v_0 \longrightarrow v_n$. Now define the inclusion $i : Q \longrightarrow U(F_C(Q))$ where each vertex and edge is sent identically. That is, vertices $v$ map to $v$ in $F_C(Q)$, and morphisms are sent identically and for each edge $e : v \longrightarrow v'$:

$$i(e : v \longrightarrow v') = (v, e, v').$$

An important observation to make is the fact that every morphism $(v_0, e_0e_1 \cdots e_{n-1}, v_n) : v_0 \longrightarrow v_n)$ in $F_C(Q)$ is a composition of length 2-morphism:

$$v_0 \xrightarrow{e_0} v_1 \xrightarrow{e_1} \cdots \xrightarrow{e_{n-1}} v_n$$

$$= \quad (v_0 \xrightarrow{e_0} v_1) \quad \circ \quad (v_1 \xrightarrow{e_1} v_2) \quad \circ \cdots \circ \quad (v_{n-1} \xrightarrow{e_{n-1}} v_n)$$

Therefore, for any graph homomorphism $\varphi : Q \longrightarrow U(\mathcal{C})$, we can create a unique functor $F : F_C(Q) \longrightarrow \mathcal{C}$ where

$$v \longmapsto \varphi(v)$$
$$(v_0, e_0e_1 \cdots e_{n-1}, v_n) : v_0 \longrightarrow v_n \longmapsto \varphi(e_0 : v_0 \longrightarrow v_1) \circ \varphi(e_1 : v_1 \longrightarrow v_2) \circ \cdots \circ \varphi(e_{n-1} : v_{n-1} \longrightarrow v_n)$$

which then gives us

$$U(F) \circ i = \varphi$$

as desired. ∎

## 2.7 Quotient Categories

The quotient category is a concept that generalizes the ideas of forming quotient groups, rings, modules, and even topological spaces. The core idea of obtaining a quotient "object" revolves around the concept of an equivalence class.

For example, in constructing the quotient group, one can go about constructing it in two different ways. One is easy, in which you simply form the concept of a coset, and then observe that nice things happen when you make cosets with normal subgroups. The hard way is to construct an equivalence relation, which *gives rise* to what we recognize as the concept of a coset, and then continuing further to create the quotient groups from normal subgroups. Both ways are equivalent, but one ignores the crucial and powerful idea of equivalence relations.

**Definition 2.7.1.** Let $\mathcal{C}$ be a locally small category. Suppose $R$ is a function which, for every pair of objects $A, B$, assigns equivalence relations $\sim_{A,B}$ on the hom set $\text{Hom}_{\mathcal{C}}(A, B)$. Then we may define the quotient category $\mathcal{C}/R$ where
**Objects.** The same objects of $\mathcal{C}$.
**Morphisms.** For any objects $A, B$ of $\mathcal{C}$, we set $\text{Hom}_{\mathcal{C}/R}(A, B) = \text{Hom}_{\mathcal{C}}(A, B)/\sim_{A,B}$.

Thus we see that morphisms between $f : A \longrightarrow B$ in $\mathcal{C}$ becomes equivalence classes $[f]$ in $\mathcal{C}/R$.

With that said, we can naturally define a canonical functor $Q : \mathcal{C} \longrightarrow \mathcal{C}/R$ where $Q$ acts identically on objects and where $Q(f : A \longrightarrow B) = [f] \in \text{Hom}_{\mathcal{C}/R}(A, B)$. This in fact defines a functor if we observe that, for a pair of composable morphisms $g, f$.

$$Q(g) \circ Q(f) = [g \circ f] = Q(g \circ f).$$

A nice property of this functor is the fact that if $f \sim f'$, then $Q(f) = Q(f')$. What is even nicer about this functor is that it has the following property.

**Proposition 2.7.2.** Let $\mathcal{C}$ be a locally small category with an equivalence relation $\sim_{A,B}$ on each set $\text{Hom}_{\mathcal{C}}(A, B)$. Then for any functor $F : \mathcal{C} \longrightarrow \mathcal{D}$ into some category $\mathcal{D}$ such that $f \sim f'$, $F(f) = F(f')$, there exists a *unique* functor $H : \mathcal{C}/R \longrightarrow \mathcal{D}$ such that $H \circ Q = F$; or, diagrammatically, such that the following diagram commutes.

$$\begin{tikzcd}
\mathcal{C} \arrow[r, "Q"] \arrow[dr, "F"'] & \mathcal{C}/R \arrow[d, dashed, "H"] \\
& \mathcal{D}
\end{tikzcd}$$

*Proof.* Observe that one functor $H : \mathcal{C}/R \longrightarrow \mathcal{D}$ that we can supply, which will have the above diagram commute, is one where $H(C) = F(C)$ on objects and where for any $[f] \in \text{Hom}_{\mathcal{C}/R}(A, B)$,

$$H([f]) = F(f)$$

where $f$ is an representative of the equivalence class $f$. Note that this is well defined since $F(f) = F(f')$ if $f \sim_{A,B} f'$; hence this will appropriately send equivalent elements to the same morphism. It is not hard to show that it's unique; one can just suppose such an $H$ exists and then demonstrate that it behaves like the functor we proposed initially. ■

---

**Example 2.7.3.**

---

## 2.8 Monoids, Groups and Groupoids in Categories

One of the most simplest, useful and yet underrated concepts in mathematics is the concept of a monoid. The reason why monoids are so useful is because they capture three main concepts: **stacking** "things" together to create another "thing," in such a way that our stacking operation is **associative**, with the additional assumption of an **identity** element which doesn't change the value. Often times in cooking up a mathematical construction, we want to maintain these three concepts because they are so familiar to our basic human nature.

Now recall the definition of a monoid.

**Definition 2.8.1.** A monoid $M$ is a set equipped with a binary operation $\cdot : M \times M \longrightarrow M$ and an identity element $e$ such that

1. For any $x, y, z \in M$, we have that $x \cdot (y \cdot z) = (x \cdot y) \cdot z$
2. For any $x \in M$, $x \cdot e = x = e \cdot x$.

It turns out that we can abstract the above definition very easily if we just resist the temptation to explicitly refer to our elements. In order to do this, we need to find a way to diagrammatically express the above axioms.

Towards that goal, rename the binary operation as $\mu : M \times M \longrightarrow M$ (for notational convenience). Then to express axiom (1), we mean that we have 3 elements $x, y, z \in M$ and there are two ways to compute them, but we want them to be the same. So lets make each different way to compute them one side of a square, which we'll say it commutes.

$$
\begin{tikzcd}
(x,y,z) \arrow[r, mapsto, "\mu\times 1"] \arrow[d, mapsto, "1\times\mu"'] & (x\cdot y, z) \arrow[d, mapsto, "\mu"] \\
(x, y\cdot z) \arrow[r, mapsto, "\mu"'] & x\cdot(y\cdot z) = (x\cdot y)\cdot z
\end{tikzcd}
\qquad
\begin{tikzcd}
M\times M\times M \arrow[r, "\mu\times 1"] \arrow[d, "1\times\mu"'] & M\times M \arrow[d, "\mu"] \\
M\times M \arrow[r, "\mu"'] & M
\end{tikzcd}
$$

The result is the diagram on the above left. Since we want this to hold for all elements in $M$, we construct the diagram more generally on the above right; this expresses our associativity axiom. Now to express the second axiom diagrammatically, we need a way to discuss the identity map. So define the map $\eta : \{\bullet\} \longrightarrow M$ where $\eta(\bullet) = e$. This is just a stupid map that picks out the identity. Then axiom (2) can be translated diagramatically to state that the bottom left diagram commutes.

$$
\begin{tikzcd}
(\bullet, m) \arrow[r, mapsto, "\eta\times 1_M"] \arrow[dr, mapsto, "\pi_M"'] & (e,m) \arrow[d, mapsto, "\mu"] & (m,e) \arrow[d, mapsto, "\mu"] & (m,\bullet) \arrow[l, mapsto, "1_M\times\eta"'] \arrow[dl, mapsto, "1_M\times\eta"] \\
 & m = e\cdot m = & m\cdot e = m &
\end{tikzcd}
\qquad
\begin{tikzcd}
\{\bullet\}\times M \arrow[r, "\eta\times 1_M"] \arrow[dr, "\pi_M"'] & M\times M \arrow[d, "\mu"] & M\times\{\bullet\} \arrow[l, "1_M\times\eta"'] \arrow[dl, "\pi_M"] \\
 & M &
\end{tikzcd}
$$

Since we want this to hold for all $m \in M$, we generalize this to create a commutative diagram as on the above right. We now have what we need to define a monoid more generally.

**Definition 2.8.2.** Let $\mathcal{C}$ be a category with cartesian products. Denote the terminal object as $T$. An object $M$ is said to be a **monoid** in $\mathcal{C}$ if there exist maps

$$\mu : M \times M \longrightarrow M \qquad \textbf{(Multiplication)}$$
$$\eta : T \longrightarrow M \qquad \textbf{(Identity)}$$

such that the diagrams below commute.

$$\begin{tikzcd}
M \times M \times M \arrow[r, "\mu\times 1"] \arrow[d, "1\times\mu"'] & M \times M \arrow[d, "\mu"] \\
M \times M \arrow[r, "\mu"'] & M
\end{tikzcd}
\qquad
\begin{tikzcd}
T \times M \arrow[r, "\eta\times 1_M"] \arrow[dr, "\pi'_M"'] & M \times M \arrow[d, "\mu"] & M \times T \arrow[l, "1_M\times\eta"'] \arrow[dl, "\pi_M"] \\
& M &
\end{tikzcd}$$

Dually, a **comonoid** is an object $C$ with maps

$$\Delta : C \longrightarrow C \times C \qquad \textbf{(Comultiplication)}$$
$$\varepsilon : C \longrightarrow T \qquad \textbf{(Identity)}$$

such that the dual diagrams commute.

$$\begin{tikzcd}
C \arrow[r, "\Delta"] \arrow[d, "\Delta"] & C \times C \arrow[d, "\Delta\times 1_C"] \\
C \times C \arrow[r, "1_C\otimes\Delta"'] & C \times C \times C
\end{tikzcd}
\qquad
\begin{tikzcd}
T \times C & C \times C \arrow[l, "\varepsilon\times 1_C"] \arrow[r, "1_C\times\varepsilon"] & C \times T \\
& C \arrow[ul, "\sigma'"] \arrow[u, "\Delta"] \arrow[ur, "\sigma"'] &
\end{tikzcd}$$

Note that we're being a little sloppy here. For example, the object $M \times M \times M$ doesn't actually exist; we have either $M \times (M \times M)$ or $(M \times M) \times M$. However, for any category with cartesian products, we always have that these two objects are isomorphic. Hence we mean either of the equivalent products when we discuss $M \times M \times M$.

---

**Example 2.8.3.** Let $k$ be a field. Consider the category $\textbf{Vect}_k$. Then a monoid in this category is an object $A$ equipped with maps

---

---

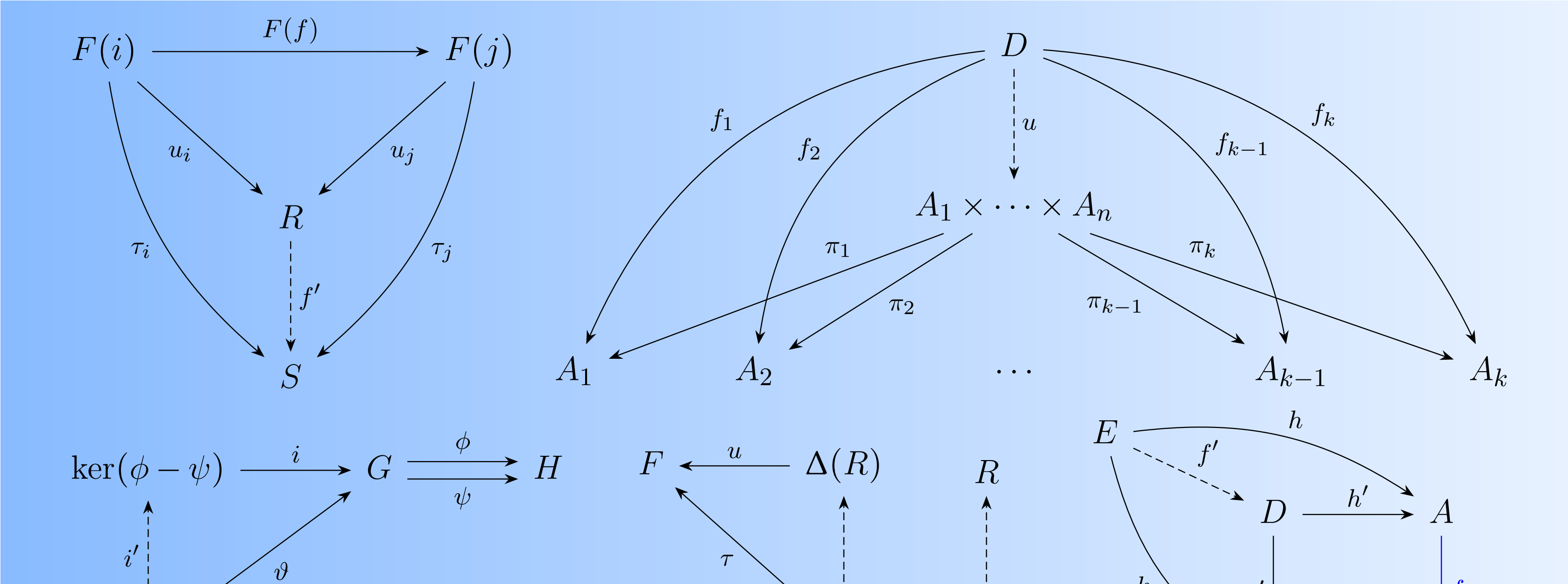
**Example 2.8.4.** Group object in the category of Top is a topological group.

---

---

**Example 2.8.5.** Monoid in the category of $R$ modules is an associative algebra.

---




# 3. Universal Constructions and Limits

## 3.1 Universal Morphisms

This chapter is probably the most important chapter in these notes. In an ideal world, this chapter would be the first chapter. However, that would pedagogically go over terribly. The discussion requires categories, functors, and natural transformations; we need the language these concepts offer to even begin to rigorously define what a universal construction even is.

But at this point, we are in fact equipped with the fundamentals. So we can now go on and define what a universal construction is, and demonstrate its prevalence in mathematics and therefore the usefulness of category theory as a convenient language to discuss these concepts.

To begin, we will motivate with a few examples.

Let $\varphi, \psi : (G, \cdot) \longrightarrow (H, +)$ be a pair of abelian[1] group homomorphisms. We now ask the question:

What is the set of all $g \in G$ such that $\varphi(g) = \psi(g)$? Is it a subgroup of $G$?

To determine this, it is equivalent to asking when $\varphi(g) - \psi(g) = 0 \implies (\varphi - \psi)(g) = 0$. Hence every such $g \in G$ lies in the kernel of $\varphi - \psi : G \longrightarrow H$, and every element in the kernel is such a desired element; so we've answered the first question. The kernel is a subgroup of $G$, so we've answered the last question. Now because this is a kernel, it has an inclusion homomorphism $i : \text{Ker}(\varphi - \psi) \longrightarrow G$. So far, our picture looks like this:

$$\begin{tikzcd} \text{Ker}(\varphi - \psi) \arrow[r, "i"] & G \arrow[r, shift left, "\varphi"] \arrow[r, shift right, "\psi"'] & H \end{tikzcd}$$

and clearly $\varphi \circ i = \psi \circ i$. Now suppose that $\sigma : K \longrightarrow G$ is another group homomorphism with the property that $\varphi \circ \sigma = \psi \circ \sigma$. Then by our previous work, this means that for each $k \in K$,

[1] The abelian-ness becomes important later.

we have that $\sigma(k) \in \text{Ker}(\varphi - \psi)$. That is,

$$\text{Im}(\sigma) \subseteq \text{Ker}(\varphi - \psi)$$

Hence instead of mapping $K$ into $G$, we can instead map $K$ into $\text{Ker}(\varphi - \psi)$, and then travel back to $G$ using $i$. So, there is **a unique** morphism $\tau : K \longrightarrow \text{Ker}(\varphi - \psi)$ such that the diagram below commutes (Prove it is unique; it shouldn't be too bad).

$$\begin{tikzcd}
\text{Ker}(\varphi - \psi) \arrow[r, "i"] & G \arrow[r, shift left, "\varphi"] \arrow[r, shift right, "\psi"'] & H \\
K \arrow[u, dashed, "\tau"] \arrow[ur, "\sigma"'] & &
\end{tikzcd}$$

What's really going on? This is an example of a universal construction. We have a "supreme" morphism $i : \text{Ker}(\varphi - \psi) \longrightarrow G$ with the property that $\varphi \circ i = \psi \circ i$. Any other morphism $\sigma : K \longrightarrow G$ with the same property that $\varphi \circ \sigma = \psi \circ \sigma$ must factor through the "supreme" morphism $i$ **in a unique way**. Uniqueness here is very important.

Now, if you haven't seen this definition before, it's going to sting a little, and you'll probably have to read it 20 times and do many, many examples (not just *look* at examples, you have to *do* some yourself) to achieve true understanding. But here we go:

**Definition 3.1.1.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor and $D$ an object of $\mathcal{D}$. Define a **universal morphism from $D$ to $F$** to be a morphism

$$u : D \longrightarrow F(C)$$

with $C \in \text{Ob}(\mathcal{C})$ and $u$ a morphism in $\mathcal{D}$ equipped with the **universal property:**

> For every morphism $f : D \longrightarrow F(C')$, there exists a **unique** morphism $h : C \longrightarrow C'$ such that the diagram below commutes.

$$\begin{tikzcd}
D \arrow[r, "u"] \arrow[dr, "f"'] & F(C) \arrow[d, dashed, "F(h)"] & C \arrow[d, dashed, "h"] \\
& F(C') & C'
\end{tikzcd}$$

The arrow $h$ is dashed, and should be read as "there exists an $h$." This is a practice that we will continue to use throughout this text.

> **Remark 3.1.2.** To the beginner, this definition will most likely make zero sense. The only way that it will make sense is to see the definition in action.

A universal arrow can also be thought of as a pair $(C, u : D \longrightarrow F(C))$. This just emphasizes that $C$ is special. This isn't really useful for us to imagine in this way right now. So you don't have to think of it as a pair, so long as you remember you're mapping to $F(C)$.

The point is that *any* arrow of the form $f : D \longrightarrow F(C')$ forces the *unique* existence of an arrow $f' : C \longrightarrow C'$ such that $F(h) \circ u = f$.

---

**Example 3.1.3.** Let $V$, $W$ be finite-dimensional vector spaces over a field $k$. Denote their bases as $\{v_1, v_2, \dots, v_n\}$ and $\{w_1, w_2, \dots, w_m\}$.

> **Q:** What does it take for a function $T : V \longrightarrow W$ to be a linear transformation?

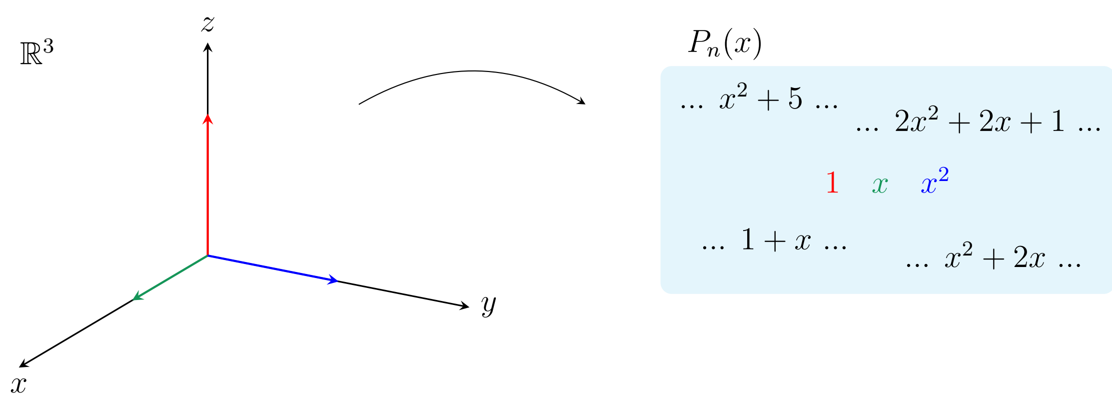
Well, suppose we have a linear transformation. Since each element of $V$ may be written as $c_1v_1 + \cdots + c_nv_n$ for $c_i \in k$, we see that

$$T(c_1v_1 + \cdots + c_nv_n) = c_1T(v_1) + \cdots + c_nT(v_n).$$

Thus we have an answer.

> **A:** To define a linear transformation $T : V \longrightarrow W$, it suffices to specify where we want $T$ to send the basis elements $v_1, \dots, v_n$.

An illustration of this fact is below.




*We can specify a linear transformation from $\mathbb{R}^3$ to the polynomial vector space $P_3(x)$ by specifying where we send the basis elements. Here, we color code where we send the basis.*

This observation helps us build our first example of universality.

Let $X$ be a (possibly infinite) set. For a field $k$, we can generate a vector space $V_x$ (Note the color-coding here corresponds to the color-coding in the definition of a universal morphism) whose basis elements are $x \in X$. Specifically,

$$V_x = \left\{ \sum_{x \in X} c_x x \;\middle|\; c_x = 0 \text{ for all but finitely many } x \right\}.$$

Now let $\mathbf{Vect}_k$ be the category of vector spaces over the field $k$. Let $U : \mathbf{Vect}_k \longrightarrow \mathbf{Set}$ be the forgetful functor which sends the vector space $V$ to the set containing *all* its elements. For any set $X$, then there is an inclusion map

$$i : X \longrightarrow U(V_x) \qquad x \mapsto x.$$

This inclusion map has the following property. Let $W$ be any vector space, and suppose that we have a function $f : X \longrightarrow U(W)$. This is kind of funny. A map $f : X \longrightarrow U(W)$ simply picks out a $w_x \in W$ for each $x \in X$. Since $X$ is a basis for $V_x$, this "picking out" defines a linear transformation $T : V \longrightarrow W$. That is, such an $f : X \longrightarrow U(W)$ allows us to define a linear transformation where for each basis element $x \in X$

$$T(x) = f(x).$$

Since we know where the basis elements go, we see that such a linear transformation is well defined. Moreover, we see that our construction makes the diagram below commute.

$$\begin{tikzcd}
X \arrow[r, "i"] \arrow[dr, "f"'] & U(V_x) \arrow[d, dashed, "U(T)"] & V_x \arrow[d, dashed, "h"] \\
& U(W) & W
\end{tikzcd}$$

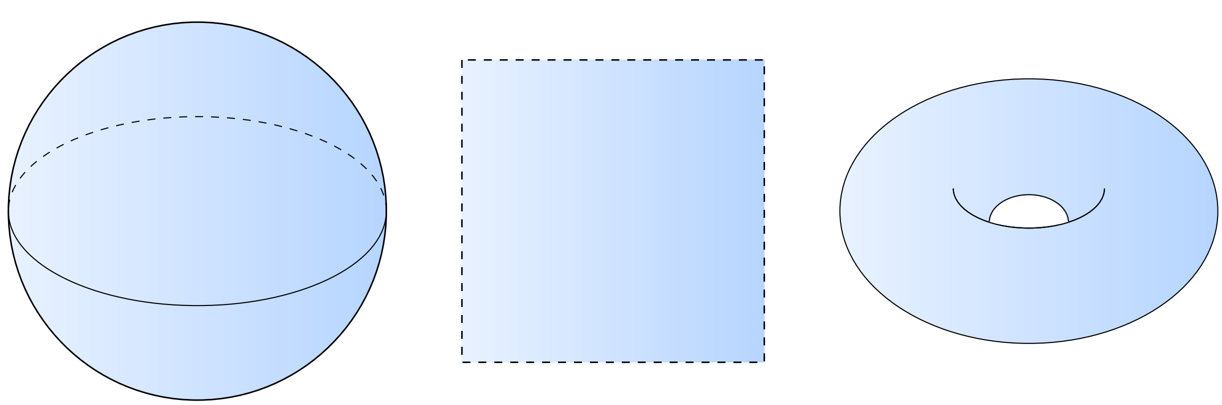
Therefore, we see that a universal morphism from $X$ to the forgetful functor $U : \mathbf{Vect}_k \longrightarrow \mathbf{Set}$ is its inclusion morphism $i : X \longrightarrow U(V_x)$ into the vector space $V_x$ generated by $X$.

---

Several key concepts in topology are secretly universal properties in disguise. This is because in some sense, the problem of universality is an optimization problem. And in elementary topology, we are often trying to optimize a given topological space with a desired property. For example, the closure of a topological space $X$ is the "largest closed set" *containing* $X$. We'll elaborate more on this.

---

**Example 3.1.4.** Let $X$ be a topological space. In topology, it is often of interest to consider a *compactification* of the space $X$. Such a story goes like this: Given $X$, we seek a compact space $X^*$ such that $X$ *embeds* as a *dense* subspace of $X^*$. In other words, we want a compact $X^*$ which has a dense subspace $S \subseteq X^*$ that is homeomorphic to $X$. We can then identify $X$ with $S$ and work within $X^*$, which is a nicer space to work inside of.



*In the middle, we have the topological space* $(0, 1) \times (0, 1)$*. As this isn't compact, we can compactify it to either (1) a sphere, by adding a point and identifying all four sides with the point, or adding sufficient points to (2) identify opposite edges to obtain a torus.*

We can, however, do even better. We can compactify $X$ into a space that is not only compact, but is also Hausdorff. The optimal compactification for this situation is the **Stone-Čech Compactification**, which is defined as follows. Given a topological space $X$, the Stone-Čech compactification is the compact, Hausdorff space $\beta X$, equipped with a dense embedding $i_X : X \longrightarrow \beta X$ such that, for any other compact, Hausdorff space $K$ equipped with a continuous map $f : X \longrightarrow K$, there exists a *unique* continuous function $\beta f : \beta X \longrightarrow K$ such that

$$\begin{tikzcd}
X \arrow[r, "i_X"] \arrow[dr, "f"'] & \beta X \arrow[d, dashed, "\beta f"] \\
& K
\end{tikzcd}$$

This universal property is what demonstrates that the Stone-Čech compactification $\beta X$ is the "most compact, Hausdorff" space we can densely embed $X$ into. However, in the language of category theory we see that this is just another example of a universal morphism. To see this, let $I : \textbf{CHaus} \longrightarrow \textbf{Top}$ be the inclusion functor from compact Hausdorff spaces into topological spaces. Then we can rewrite the diagram as

$$\begin{tikzcd}
X \arrow[r, "i_X"] \arrow[dr, "f"'] & I(\beta X) \arrow[d, dashed, "I(\beta f)"] \\
& I(K)
\end{tikzcd}$$

Of course, in practice, we'd never actually write it like this; but this is just for us to be able to see that the dense embedding $i_X : X \longrightarrow \beta X$ is universal from $X$ to the the inclusion functor $I : \textbf{CHaus} \longrightarrow \textbf{Top}$, so that the Stone-Čech compactification is truly an example of a universal morphism.

---

**Example 3.1.5.** Consider the free monoid functor $F : \textbf{Set} \longrightarrow \textbf{Mon}$ which sends a set $X$ to the free monoid generated by $X$. Specifically,

$$F(X) = \{x_1 x_2 \dots x_n \mid x_i \in X\} \cup \{e\}$$

The set consists of all strings using elements of $X$, and an identity $e$; the monoid product is concatenation.

Suppose I have a monoid homomorphism $\varphi : F(X) \longrightarrow M$ where $M$ is another monoid. Then for any two $x_1, x_2 \in X$, we have that $\varphi(x_1 x_2) = \varphi(x_1)\varphi(x_2)$. More generally, for any $x_1 \cdots x_n \in F(X)$, we have that

$$\varphi(x_1 \cdots x_n) = \varphi(x_1) \cdots \varphi(x_n).$$

We thus see the following: To define a monoid homomorphism, we just need to know where to send every individual $x \in X$. This is achieved by defining a set function $\varphi_0 : X \longrightarrow U(M)$, and by setting $\varphi(x) = \varphi_0(x)$. This makes the diagram below commutative.

$$\begin{tikzcd} X \arrow[r, "i_X"] \arrow[dr, "\varphi_0"'] & U(F(X)) \arrow[d, dashed, "\varphi"] \\ & U(M) \end{tikzcd} \qquad \begin{tikzcd} x \arrow[r, mapsto] \arrow[dr, mapsto] & i_X(x) = x \arrow[d, mapsto] \\ & \varphi_0(x) = \varphi(x) \end{tikzcd}$$

We thus see that $(X, i_X : X \longrightarrow F(X))$ is universal from $X$ to $U : \mathbf{Mon} \longrightarrow \mathbf{Set}$.

---

---

**Example 3.1.6.** Let $(R, +, \cdot)$ be a ring and $k$ a field. Suppose further that $R$ is a $k$-algebra. Then for any set $X = \{x_1, \dots, x_n\}$ of indeterminates, we can create a **free algebra generated by** $X$, denoted as $k\{X\}$. One can show that this defines a functor

$$F : \mathbf{Set} \longrightarrow \mathbf{Alg}_k$$

mapping sets $X$ it $k\{X\}$ and functions $f : X \longrightarrow Y$ to the $k$-algebra morphism $\varphi : k\{X\} \longrightarrow k\{Y\}$ where $\varphi$ is defined linearly by its action on the basis elements sending each $x \longrightarrow f(x)$. On the other hand, note that we can also create a forgetful functor

$$U : \mathbf{Alg}_k \longrightarrow \mathbf{Set}$$

which simply reinterprets each $k$-algebra as a set and each $k$-algebraic morphism as a function.

Now consider a mapping $f : X \longrightarrow U(R)$ in **Set**. Because we also have a mapping $i : X \longrightarrow U(F(X))$, which acts an inclusion function, we see that we can create a mapping $h : F(X) \longrightarrow A$ such that the diagram below commutes.

$$\begin{tikzcd} X \arrow[r, "i"] \arrow[dr, "g"'] & U(F(X)) \arrow[d, dashed, "U(h)"] \\ & U(A) \end{tikzcd} \qquad \begin{tikzcd} F(X) \arrow[d, dashed, "h"] \\ A \end{tikzcd}$$

The way we do this is we defined $h : F(X) \longrightarrow A$ to act linearly on the basis elements, sending $x \mapsto g(x)$. This defines a $k$-algebraic morphism and makes the above diagram commute. In this case, we say that $(F(X), i : X \longrightarrow U(F(X)))$ is universal from $X$ to the forgetful functor $U : \mathbf{Alg}_k \longrightarrow \mathbf{Set}$.

---

When we discuss a universal morphism from $D$ to $F : \mathcal{C} \longrightarrow \mathcal{D}$, we are particularly discussing a morphism $u : D \longrightarrow F(C)$ and a special object $C$. Hence, we can actually write a universal morphism as a pair $(C, u : D \longrightarrow F(C))$. Does this look familiar? This is an object of the category $(D \downarrow F)$! Hence, universal morphisms can actually be thought of as elements in a comma category. Under this intepretation, what does the universal property translate to? The next proposition answers our question.

**Proposition 3.1.7.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor. A morphism $u : D \longrightarrow F(C)$ is universal from $D$ to $F$ if and only if $(C, u : D \longrightarrow F(C))$ is an initial object of the comma category $(D \downarrow F)$

So, as we will see, the universal property of a universal morphism $u : D \longrightarrow F(C)$ translates to $(C, u : D \longrightarrow F(C))$ being an initial object in some comma category.

*Proof.* Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor, and $D$ an object of $\mathcal{D}$. Recall that the category $(D \downarrow \mathcal{D})$ is the category where

**Objects.** Pairs $(C, f : D \longrightarrow F(C))$ with $C \in \mathcal{C}$ and $f : D \longrightarrow \mathcal{D}$ a morphism in $\mathcal{D}$.

**Morphisms.** Morphisms between two objects $(C, f : D \longrightarrow F(C))$ and $(C', f : D \longrightarrow F(C'))$ are given by morphisms $h : C \longrightarrow C'$ such that the diagram below commutes.

$$\begin{tikzcd} & D \arrow[dl, "f"'] \arrow[dr, "f'"] & \\ F(C) \arrow[rr, "F(h)"'] & & F(C') \end{tikzcd}$$

Suppose $(A, u : D \longrightarrow F(A))$ is an initial object in $(D \downarrow F)$. Then for every other pair $(A, f : D \longrightarrow F(A'))$, there exists a unique morphism $h : A \longrightarrow A'$ such that the diagram on the bottom left commutes.

$$\begin{tikzcd} & D \arrow[dl, "u"'] \arrow[dr, "f"] & \\ F(A) \arrow[rr, dashed, "F(h)"'] & & F(A') \end{tikzcd} = \begin{tikzcd} D \arrow[r, "u"] \arrow[dr, "f"'] & F(A) \arrow[d, dashed, "F(h)"] \\ & F(A') \end{tikzcd} \qquad \begin{tikzcd} A \arrow[d, dashed, "h"] \\ A' \end{tikzcd}$$

However, if we rearrange this we see that this is just the universal property in disguise! Conversely, any pair $(A, f : A \longrightarrow F(A))$ being a universal morphism can be demonstrated to be an initial object in $(D, \downarrow F)$ by reversing the above proof. ∎

Now, we didn't do this just for fun. The interpretation of a universal morphism as an initial object of a comma category theory will serve to be very useful, just not now. As of now it does not really grant us much. But when we are deep into the chapter on Limits, this intrepretation will become useful.

One thing that the interpretation does grant us for now is the following theorem, which requires essentially no proof if we understand a universal morphism is an initial object of a comma category. This theorem explains ultimately why we care about universal morphisms; they're like categorical invariants!

**Theorem 3.1.8.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor and $D \in \mathcal{D}$. Suppose $u : D \longrightarrow F(C)$ is universal from $D$ to $F$ for some object $C \in \mathcal{C}$. If $u' : D \longrightarrow F(C')$ is also universal from $D$ to $F$, then $C \cong C'$.

*Proof.* Universal morphisms $u : D \longrightarrow F(C)$ are initial objects in the comma category $(D \downarrow F)$, and initial objects are always unique up to isomorphism. Hence $(C, u : D \longrightarrow F(C))$ with the universal property is unique. ∎

However, the direct proof, where we do not use the interpretation of a comma category, is left as an exercise. It's actually very important to see and understand the direct proof.

As with most constructions within category theory, there is a dual construction. That, is there is another form of universality which is equally as important as the one we originally introduced. So, in general, there are two forms of universality.

**Definition 3.1.9.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor and $C$ an object of $\mathcal{C}$. A **universal arrow from $F$ to $C$** is a morphism

$$v : F(C) \longrightarrow D$$

equipped with the **universal property**:

> For every $f : F(C') \longrightarrow D$, there exists a a unique morphism $h' : C' \longrightarrow C$ such that the diagram below commutes.

$$\begin{tikzcd} F(C) \arrow[r, "v"] & D & & C \\ & & & \\ F(C') \arrow[uu, dashed, "F(f')"] \arrow[uur, "f"'] & & & C' \arrow[uu, dashed, "h"'] \end{tikzcd}$$

Note that this is basically the previous definition of a universal arrow from an object to a functor, except the direction of the arrows have been flipped. This is why we called this the "dual" definition of the previous one. This motivates the following statement which requires no effort to prove.

**Proposition 3.1.10.** Let $\mathcal{C}$ be a category and $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor. If $\mathcal{C}$ has a universal morphism from $D$ to $F$, then $\mathcal{C}^{\text{op}}$ has a universal morphism from $F$ to $D$.

So we see that the two notions of unviversality we've introduced really are dual concepts. Both are equally important, and we will see that they both arise as very deep concepts in mathematics. Not just in the examples we've provided, but in deeper pure category theory.

Anyways, we can repeat the propositions we worked on.

**Proposition 3.1.11.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor. A morphism $u : F(C) \longrightarrow D$ is universal from $F$ to $D$ if $(C, u : F(C) \longrightarrow D)$ is a terminal object of the comma category $(F \downarrow D)$.

This is left as an exercise, and should be similar to our proof from before. And as before, we get our second important theorem:

**Theorem 3.1.12.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor, with $u : F(C) \longrightarrow D$ universal from $F$ to $D$. Then if $u' : F(C') \longrightarrow D$ is also universal from $F$ to $D$, then $C \cong C'$.

*Proof.* Universal morphisms from $F$ to $D$ are terminal objects in a comma category, and terminal objects are always unique up to isomorphism. ∎

The direct proof is also an exercise.

## Exercises

**1.** Prove Theorem 3.1.8 directly, and dualize your proof to prove Theorem 3.1.12 directly.

**2.** Prove Proposition 3.1.11.

**3.** For each ring $R$, we may construct the single-variable polynomial ring $R[x]$. This process defines a functor **Poly** : **Ring** $\longrightarrow$ **Ring**.

Show that for each ring $R$, the inclusion ring homomorphism $i : R \longrightarrow R[X]$ is a universal morphism from $R$ to **Poly**.

**4.** Let $X$ and $Y$ be two sets, and consider their product $X \times Y$. Recall that with any product, we have "projection maps" $\pi_1 : X \times Y \longrightarrow X$ and $\pi_2 : X \times Y \longrightarrow Y$ where $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$.

*i.* Suppose we have functions $f : Z \longrightarrow X$ and $g : Z \longrightarrow Y$. Show how this gives us a map $h : Z \longrightarrow X \times Y$, and show that this map is unique (to the pair $f$ and $g$).

*ii.* Using your map $h : Z \longrightarrow X \times Y$, show that the diagram on the left commutes, and that the diagram on the right is equivalent.

$$\begin{tikzcd} & Z \arrow[dl, "f"'] \arrow[d, dashed, "h"] \arrow[dr, "g"] & \\ X & X \times Y \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & Y \end{tikzcd} \qquad \begin{tikzcd} (X \times Y, X \times Y) \arrow[r, "(\pi_1,\pi_2)"] & (X, Y) \\ (Z, Z) \arrow[u, dashed, "(h,h)"] \arrow[ur, "(f,g)"'] & \end{tikzcd}$$

To be clear, the diagram on the right is in the category **Set** $\times$ **Set**.

*iii.* Let $\Delta$ : **Set** $\longrightarrow$ **Set** $\times$ **Set** be the "copy functor" which sends $X \mapsto (X, X)$. Then the above diagram translates to

$$\begin{tikzcd}
\Delta(X \times Y) \arrow[r, "(\pi_1,\pi_2)"] & (X,Y) \\
\Delta(Z) \arrow[u, dashed, "\Delta(h)"] \arrow[ur, "(f,g)"'] &
\end{tikzcd}$$

Deduce how the product $(\pi_1, \pi_2) : \Delta(X \times Y) \longrightarrow (X, Y)$ is universal from $(X, Y)$ to $\Delta$. This is an important fact that we'll build upon later.

**4.** Let $X$ and $Y$ be two sets, and consider the coproduct

$$X \amalg Y = \{(x, 1), (y, 2) \mid x \in X, y \in Y\}^2$$

Recall that with any coproduct, we'll have "injection maps" $i_1 : X \longrightarrow X \amalg Y$ and $i_2 : Y \longrightarrow X \amalg Y$ where $i_1(x) = (x, 1)$ and $i_2(y) = (y, 2)$. Repeat (*i*-*iii*) as in the previous exercise to demonstrate that $(i_1, i_2) : (X, Y) \longrightarrow \Delta(X \amalg Y)$ is universal from $\Delta$ to $(X, Y)$.

[2]Note that I arbitrarily chose the numbers 1 and 2. I could have put anything I wanted. For a coproduct, we just need to create two separate tuples that contain $x$ values and $y$-values. Hence 1 and 2 work perfectly fine.

## 3.2 Representable Functors and Yoneda's Lemma

This is probably the most important section out of these entire set of notes. The propositions proved here will allows us to perform slick proofs of interesting results later on. We will also use results that are left as exercises for the reader (since it is important for the reader to do them). Now before we introduce the Yoneda lemma, we prove some propositions concerning the concept of universality.

**Proposition 3.2.1.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor. Then a pair $(R, u : D \longrightarrow F(R))$ is universal from $D$ to $F$ if and only if for each $C \in \mathcal{C}$ we have the natural bijection

$$\mathrm{Hom}_{\mathcal{C}}(R, C) \cong \mathrm{Hom}_{\mathcal{D}}(D, F(C)).$$

That is, any isomorphism, natural in $C$ as above, is determined by a unique morphism $u : D \longrightarrow F(R)$ so that $(R, u)$ is a universal arrow from $D$ to $F$.

*Proof.* Suppose that $u : D \longrightarrow F(R)$ is a universal morphism from $D$ to $F$. Then by definition, we have the relation

$$\begin{tikzcd}
D \arrow[r, "u"] \arrow[dr, "h"'] & F(R) \arrow[d, dashed, "F(f)"] & R \arrow[d, dashed, "f"] \\
& F(C) & C
\end{tikzcd}$$

Each $h : D \longrightarrow F(C)$ uniquely corresponds to a morphism $f : R \longrightarrow C$, while conversely, any $f : R \longrightarrow C$ can be precomposed with $u$ to obtain a morphism $F(f) \circ u : D \longrightarrow F(C)$. Hence we see the we have a bijective correspondence

$$\mathrm{Hom}_{\mathcal{D}}(R, C) \cong \mathrm{Hom}_{\mathcal{C}}(D, F(C)).$$

Now to demonstrate naturality, we consider a morphism $k : C \longrightarrow C'$ and we check that the diagram below commutes.

$$\begin{tikzcd}
\mathrm{Hom}_{\mathcal{D}}(R, C) \arrow[r, "\sim"] \arrow[d, "k\circ(-)"'] & \mathrm{Hom}_{\mathcal{C}}(D, F(C)) \arrow[d, "F(k)\circ(-)"] \\
\mathrm{Hom}_{\mathcal{D}}(R, C') \arrow[r, "\sim"] & \mathrm{Hom}_{\mathcal{C}}(D, F(C'))
\end{tikzcd}$$

$$\begin{tikzcd}
D \arrow[r, "u"] \arrow[dr, "h"] \arrow[ddr, "F(k)\circ F(f)\circ u"'] & F(R) \arrow[d, dashed, "F(f)"] \arrow[dd, bend left, "F(k\circ f)"] & R \arrow[d, dashed, "f"] \arrow[dd, bend left, "k\circ f"] \\
& F(C) \arrow[d, "F(k)"] & C \arrow[d, "k"] \\
& F(C') & C'
\end{tikzcd}$$

- Beginning with a morphism $f : R \longrightarrow C$, we travel right to obtain the morphism $F(f) \circ u$. Going down, we obtain the morphism $F(k) \circ (F(f) \circ u)$.
- Consider the same morphism $f : R \longrightarrow C$. If we instead first traveled down, we'd obtain the morphism $k \circ f$. Traveling right would then send us to the morphism $F(k \circ f) \circ u$.

However, it is certainly the case that

$$F(k) \circ (F(f) \circ u) = F(k \circ f) \circ u$$

so that these paths are equivalent. The proof could also be given immediately by considering the diagram on the left, which is supplied here to give a better understanding of what's going on.

To prove the other direction, suppose that we have such a natural bijection given by some $\varphi$.

$$\varphi_C : \text{Hom}_{\mathcal{D}}(R, C) \xrightarrow{\sim} \text{Hom}_{\mathcal{C}}(D, F(C))$$

Then in particular we have that $\text{Hom}_{\mathcal{D}}(R, R) \cong \text{Hom}_{\mathcal{C}}(D, F(R))$. Consider $\varphi(1_R) : D \longrightarrow F(R)$; we denote this special morphism as $u : D \longrightarrow F(R)$.

Now for any $f : R \longrightarrow C$, the diagram on the bottom left commutes by naturality; however, we are more interested in following the element $1_R \in \text{Hom}_{\mathcal{D}}(R, R)$.

$$\begin{tikzcd}
\text{Hom}_{\mathcal{D}}(R, R) \arrow[r, "\sim"] \arrow[d, "f\circ(-)"'] & \text{Hom}_{\mathcal{C}}(D, F(R)) \arrow[d, "F(f)\circ(-)"] \\
\text{Hom}_{\mathcal{D}}(R, C) \arrow[r, "\sim"] & \text{Hom}_{\mathcal{C}}(D, F(C))
\end{tikzcd}
\qquad
\begin{tikzcd}
1_R \arrow[r, mapsto] \arrow[d, mapsto] & u : D \longrightarrow F(R) \arrow[d, mapsto] \\
f : R \longrightarrow C \arrow[r, mapsto] & \varphi(f) = F(f) \circ u
\end{tikzcd}$$

We see that any such $\varphi$ must act on $\text{Hom}_{\mathcal{D}}(R, C)$ by bijectively send $f : R \longrightarrow C$ to $F(f) \circ u$. What this means is that any $h \in \text{Hom}_{\mathcal{C}}(D, F(C))$ corresponds uniquely to some $f : R \longrightarrow C$ such that $h = F(f) \circ u$, which is exactly the definition for $u : D \longrightarrow F(R)$ to be universal from $D$ to $F$. This completes the proof. ∎

In the proof we demonstrated above, we did something weird. That is, we discussed this so-called natural isomorphism

$$\varphi_C : \text{Hom}_{\mathcal{D}}(R, C) \longrightarrow \text{Hom}_{\mathcal{C}}(D, F(C)).$$

However, at this point we've only really seen natural isomorphisms *between functors*. Does this mean what we really had was a natural transformation between two functors? The answer is yes; the proof inadvertently derived the natural isomorphism

$$\varphi : \text{Hom}_{\mathcal{D}}(R, -) \longrightarrow \text{Hom}_{\mathcal{C}}(D, F(-))$$

which, by the proposition above, exists only when we have a universal morphism $u : D \longrightarrow F(R)$ from $D$ to $F$. For such functors, we call them *representable*.

**Definition 3.2.2.** Let $\mathcal{C}$ have small hom-sets. We say a functor $K : \mathcal{C} \longrightarrow \textbf{Set}$ is representable when there exists an object $R$ and a natural isomorphism

$$\psi : \text{Hom}_{\mathcal{D}}(R, -) \longrightarrow K.$$

The object $R$ here is said to be the **representing object** for $K$.

---

**Example 3.2.3.** Consider the forgetful functor $U : \mathbf{Grp} \longrightarrow \mathbf{Set}$. One way to describe this functor is simply with words: each group $G$ is sent to its underlying set in **Set**. Another approach is to literally express the groups in terms of its elements, for this then tells us where it is sent in **Set**. A simple way to do this is to consider the maps

$$\text{Hom}_{\mathbf{Grp}}(\mathbb{Z}, G) = \{\text{Group homomorphisms } \varphi : \mathbb{Z} \longrightarrow G\}.$$

This works since each such map $\varphi : \mathbb{Z} \longrightarrow G$ firstly picks out some element $a$ so that $\varphi(1) = a$. As this is a group homomorphism we then see that $\varphi(n) = a^n$. Hence the collection of all these maps picks out all of the elements of $G$, so that we can say

$$U(G) \cong \text{Hom}_{\mathbf{Grp}}(\mathbb{Z}, G).$$

We use an isomorphism since an equality is not exactly correct; we just know that the two sets are going to have the same cardinality, and hence be isomorphic in **Set**. Now, what this in the end means is that the forgetful functor is a representable, since we have that

$$U : \mathbf{Grp} \longrightarrow \mathbf{Set} \cong \text{Hom}(\mathbb{Z}, -) : \mathbf{Grp} \longrightarrow \mathbf{Set}.$$

This construction works due to the key property of the group homomorphism, so that this can be repeated for **Ring**, $R$-**Mod**, etc. Hence many forgetful functors are representable functors. We will see in Chapter 5 what this really means.

---

---

**Example 3.2.4.** Let $(R, +, \cdot)$ be a ring and $(k, +, \cdot)$ a field. Suppose further that $R$ is $k$-algebra. Recall that we can create the affine $n$-space of $R$

$$A^n(R) = \{(x_1, \dots, x_n) \mid x_i \in R\}.$$

Now suppose $\varphi : R \longrightarrow S$ is a morphism of $k$-algebras. Then this induces a mapping

$$A^n(\varphi) : A^n(R) \longrightarrow A^n(S) \qquad (r_1, \cdots, r_n) \mapsto (\varphi(r_1), \dots, \varphi(r_n)).$$

What we can realize now is that we have a functor on our hands (by of course verifying the other necessary properties) between $\mathbf{Alg}_k$ and **Set**.

$$A^n : \mathbf{Alg}_k \longrightarrow \mathbf{Set}.$$

Now recall from Example 2.3.1 that if $F : \mathbf{Set} \longrightarrow \mathbf{Alg}_k$ is the free functor assigning $X \mapsto k\{X\}$, the free algebra, and $U : \mathbf{Alg}_k \longrightarrow \mathbf{Set}$ is the forgetful functor, then for each set $X$ we have a universal morphism $(F(X), i : X \longrightarrow U(F(X)))$ from $X$ to the forgetful functor $U$. By Proposition **??**, we thus have the isomorphism

$$\mathrm{Hom}_{\mathbf{Alg}_k}(F(X), R) \cong \mathrm{Hom}_{\mathbf{Set}}(X, U(R)).$$

natural for all $R \in \mathbf{Alg}_k$. However, notice that if $X = \{x_1, \dots, x_n\}$, $\mathrm{Hom}_{\mathbf{Set}}(X, U(R))$ is nothing more than the set of all functions which pick out $n$ elements of $R$. In other words,

$$\mathrm{Hom}_{\mathbf{Set}}(X, U(R)) \cong A^n(R).$$

One can verify the naturality of the above bijection (I won't it's not too bad). Therefore we have that

$$\mathrm{Hom}_{\mathbf{Alg}_k}(F(X), R) \cong A^n(R) \implies \mathrm{Hom}_{\mathbf{Alg}_k}(K\{X\}, R) \cong A^n(R).$$

so that we have a natural isomorphism between functors

$$\mathrm{Hom}_{\mathbf{Alg}_k}(K\{X\}, -) \cong A^n(-).$$

What this then means is that $A^n(-)$ is a representable functor.

---

---

**Example 3.2.5.** Let $X$ be a topological space. Recall from Example ?? that we can consider the set $\mathrm{Path}(X)$ consisting of all paths in the topological space $X$. If we recall that a path in $X$ can be represented by a continuous function $f : [0, 1] \longrightarrow X$, we see that

$$\mathrm{Path}(X) = \{f : [0, 1] \longrightarrow X \mid f \text{ is continuous}\} = \mathrm{Hom}_{\mathbf{Top}}([0, 1], X).$$

Hence we see that Path : $\mathbf{Top} \longrightarrow \mathbf{Set}$ is a functor; moreover, it is clearly representable since $\mathrm{Path}(-) = \mathrm{Hom}_{\mathbf{Top}}([0, 1], -)$.

This example, however, can be taken even further: What about $n$-dimensional "paths?" To generalize this we can use simplicies. Denote $\Delta^n$ as the $n$-simplex. Then we can establish the family of functors

$$\mathrm{Hom}_{\mathbf{Top}}(\Delta^n, -) : \mathbf{Top} \longrightarrow \mathbf{Set}$$

which map simplicies to topological spaces; such continuous functions provide the foundation for singularly homology theory, and each functor above is representable . Note that we get back Path when $n = 1$.

---

As we have just seen, representable functors not only occur very frequently but they also arise naturally to yield consturctions which we actually care about.

A natural question to ask at this point is the following: When exactly do we have a representable functor on our hands? The next proposition answers that question.

**Proposition 3.2.6.** Let $\mathcal{C}$ be a locally small category, and suppose $K : \mathcal{C} \longrightarrow \textbf{Set}$ is a functor. Then $K$ is a representable functor (with representing object $R$) if and only if $(R, u : \{\bullet\} \longrightarrow K(R))$ is universal from $\{\bullet\}$ to $K$ for some object $R \in \mathcal{C}$.

Note here that $\{\bullet\}$ is the one-point set whose single element is denoted as $\bullet$.

*Proof.* The forward direction is similar to Example 3.2, while the backwards direction is similar to the proof of Proposition **??**.

First let's interpret what it means for $u : \{\bullet\} \longrightarrow K(R)$ to be universal. This means that for any other $f : \{\bullet\} \longrightarrow K(C')$, there exists a unique morphism $h : R \longrightarrow C'$ such that the diagram below commutes.

$$\begin{tikzcd}
\{\bullet\} \arrow[r, "u"] \arrow[dr, "f"'] & K(R) \arrow[d, dashed, "K(h)"] & & R \arrow[d, dashed, "h"] \\
& K(C') & & C'
\end{tikzcd}$$

By Proposition **??** we also have the natural bijection

$$\mathrm{Hom}_{\mathcal{C}}(R, C) \cong \mathrm{Hom}_{\textbf{Set}}(\{\bullet\}, K(C))$$

which is enough to establish a natural isomorphism $\varphi : \mathrm{Hom}_{\mathcal{C}}(R, -) \cong \mathrm{Hom}_{\textbf{Set}}(\{\bullet\}, K(-))$.

Now observe that for a given $C'$, each $f \in \mathrm{Hom}_{\textbf{Set}}(\{\bullet\}, K(-))$ is just a function $f : \{\bullet\} \longrightarrow K(R)$. Thus, each function can be represented uniquely by an element $c \in K(C)$, which establishes the bijection

$$\mathrm{Hom}_{\textbf{Set}}(\{\bullet\}, K(C)) \cong K(C)$$

for each $C$. In fact, it's not difficult to show that this bijection is natural. Therefore we see that we can connect our natural bijections together

$$\mathrm{Hom}_{\mathcal{C}}(R, -) \cong \mathrm{Hom}_{\textbf{Set}}(\{\bullet\}, K(-)) \cong K(-)$$

which demonstrates that $K : \mathcal{C} \longrightarrow \textbf{Set}$ is a representable functor.

Conversely, suppose that $K : \mathcal{C} \longrightarrow \textbf{Set}$ is representable. Specifically, suppose $\varphi : \mathrm{Hom}_{\mathcal{C}}(R, -) \xrightarrow{\sim} K(-)$ is our natural isomorphism between the functors. Then in particular, for any $h : R \longrightarrow C$, naturality guarantees that the following diagram commutes.

$$\begin{tikzcd}
\mathrm{Hom}_{\mathcal{C}}(R, R) \arrow[r, "\sim"] \arrow[d, "h\circ(-)"'] & K(R) \arrow[d, "K(h)\circ(-)"] & 1_R \arrow[rr, mapsto] \arrow[d, mapsto] & & \varphi(1_R) \arrow[d, mapsto] \\
\mathrm{Hom}_{\mathcal{C}}(R, C) \arrow[r, "\sim"] & K(C) & h : R \longrightarrow C \arrow[rr, mapsto] & & \varphi(h) = K(h)\big(\varphi(1_R)\big).
\end{tikzcd}$$

Now take a step back; define the morphism $u : \{\bullet\} \longrightarrow K(R)$ where $u(\bullet) = \varphi(1_R)$, and suppose $f : \{\bullet\} \longrightarrow K(C)$ is some morphism. Then because $\varphi : \text{Hom}_{\mathcal{C}}(R, C) \longrightarrow K(C)$ is a bijection, this means that $f(\bullet) = \varphi(h : R \longrightarrow C)$ for some **unique** morphism $h : R \longrightarrow C$. In particular, the above diagram tells us that

$$K(h)\Big(\varphi(1_R)\Big) = \varphi(h) \implies K(h)(u(\bullet)) = f(\bullet).$$

In other words, we have that given any $f : \{\bullet\} \longrightarrow K(C)$, there exists a unique $h : R \longrightarrow C$ such that the diagram commutes.

$$\begin{tikzcd}
\{\bullet\} \arrow[r, "u"] \arrow[dr, "f"'] & K(R) \arrow[d, dashed, "K(h)"] \\
& K(C')
\end{tikzcd}$$

Therefore, the fact that $K$ is representable gives rise to a $u : \{\bullet\} \longrightarrow K(R)$ which is universal, which is what we set out to show. ■

We are now ready to introduce the well-known lemma due to Nobuo Yoneda. The Yoneda lemma is simply a convenient result that occurs when one encounters situations with the functors $\text{Hom}_{\mathcal{C}}(R, -) : \mathcal{C} \longrightarrow \textbf{Set}$. While this might not seem that relevant, it applicability expands when we combine the result with our previous work on representable functors in this section.

**Theorem 3.2.7. (Yoneda "Lemma")** Let $K : \mathcal{C} \longrightarrow \textbf{Set}$ be a functor. Then for every object $R$ of $\mathcal{C}$, we have that

$$\text{Hom}_{\textbf{Set}^{\mathcal{C}}}\Big(\text{Hom}_{\mathcal{C}}(R, -), K\Big) \cong K(R) \implies \text{Nat}(\text{Hom}_{\mathcal{C}}(R, -), K) \cong K(R)$$

where $\text{Nat}(F, G)$ denotes the set of all natural transformations between functors $F, G$.

*Proof.* To demonstrate bijectivity, we construct two maps from each set and demonstrate that they are inverses.

Suppose we have a natural transformation $\eta : \text{Hom}_{\mathcal{C}}(R, -) \longrightarrow K$. Then for every $C \in \mathcal{C}$, the diagram below on the left commutes.

$$\begin{tikzcd}
R \arrow[d, "f"] & \text{Hom}_{\mathcal{C}}(R,R) \arrow[r, "\eta_R"] \arrow[d, "f\circ(-)"'] & K(R) \arrow[d, "K(f)"] & 1_A \arrow[r, mapsto] \arrow[d, mapsto] & \eta_R(1_R) = u \arrow[d, mapsto] \\
C & \text{Hom}_{\mathcal{C}}(R,C) \arrow[r, "\eta_C"'] & K(C) & f \arrow[r, mapsto] & \eta_C(f) = K(f)(u)
\end{tikzcd}$$

With this diagram, we can follow what happens to the identity morphism $1_R \in \text{Hom}_{\mathcal{C}}(R, R)$. As above, denote $\eta_R(1_R) = u \in K(R)$. The commutativity of the diagram above then tells us that

$$\eta_C(f : R \longrightarrow C) = K(f)(u).$$

This is great! This tells us the exact formula for every $\eta \in \text{Nat}(\text{Hom}_{\mathcal{C}}(R, -), K)$. Moreover, each formula is uniquely determined by some $u \in K(R)$. This then motivates us to construct the mapping

$$y : \text{Nat}(\text{Hom}_{\mathcal{C}}(R, -), K) \longrightarrow K(R) \qquad \eta \mapsto u$$

where $u$ is the unique member of $K(R)$ such that $\eta_C(f : R \longrightarrow C) = K(f)(u)$.

Now consider any arbitrary member $r \in K(R)$. For each $C \in \mathcal{C}$, construct the mapping

$$\varepsilon_C : \text{Hom}_{\mathcal{C}}(R, C) \longrightarrow K(R) \qquad \varepsilon_C(f : R \longrightarrow C) = K(f)(r)$$

This defines a natural transformation, so that what we've constructed is a mapping

$$y' : K(R) \longrightarrow \text{Nat}(\text{Hom}_{\mathcal{C}}(R, -), K) \qquad r \mapsto \varepsilon_C$$

where $\varepsilon_C(f : R \longrightarrow C) = K(f)(u)$.

Now given any $\eta \in \text{Nat}(\text{Hom}_{\mathcal{C}}(R, -), K)$ we clearly have that $y' \circ y(\eta) = \eta$ and for any $r \in K(r)$ we have that $y \circ y'(r) = r$. Hence we have a bijection between sets, so we may conclude that

$$\text{Nat}(\text{Hom}_{\mathcal{C}}(R, -), K) \cong K(R)$$

as desired. ∎

---

**Example 3.2.8.** As the Yoneda lemma is a bit mysterious when one first encounters it, we can perform a simple sanity check as follows. For any category $\mathcal{C}$, consider the objects $A, B \in \mathcal{C}$, which we can use to build the functors $\text{Hom}(A, -), \text{Hom}(B, -) : \mathcal{C} \longrightarrow \textbf{Set}$. What is a natural transformation $\eta : \text{Hom}(A, -) \longrightarrow \text{Hom}(B, -)$? It is a family of *functions*, indexed by all objects in $\mathcal{C}$, such that for each $f : C \longrightarrow D$ the diagram below commutes.

$$\begin{tikzcd}
C \arrow[d, "f"] & \text{Hom}(A, C) \arrow[r, "\eta_C"] \arrow[d, "f\circ(-)"'] & \text{Hom}(B, C) \arrow[d, "f\circ(-)"] & k : A \longrightarrow C \arrow[r, mapsto] \arrow[d, mapsto] & \eta_C(k) : B \longrightarrow C \arrow[d, mapsto] \\
D & \text{Hom}(A, D) \arrow[r, "\eta_D"'] & \text{Hom}(B, D) & f \circ k : A \longrightarrow D \arrow[r, mapsto] & \eta_D(f \circ k) = f \circ \eta_C(k).
\end{tikzcd}$$

We see that these functions must satisfy the property outlined in yellow for all $C, D$. So what functions do this? An immediate source of such functions that assemble into natural transformations which we seek arise when we take any $\varphi \in \text{Hom}(B, A)$ and set each $\eta_C$ : $\text{Hom}(A, C) \longrightarrow \text{Hom}(B, C)$ equal to

$$(-) \circ \varphi : \text{Hom}(A, C) \longrightarrow \text{Hom}(B, C)$$

for each $C \in \mathcal{C}$. This clearly checks out since we have that, for any $f : C \longrightarrow D$ and $k : A \longrightarrow C$,

$$(f \circ k) \circ \varphi = f \circ (k \circ \varphi).$$

The question now is: Is every natural transformation derived from some $\varphi \in \text{Hom}(B, A)$? We know that the answer is yes! This is an exercise in Section 1.9. The work of that exercise is proving this; however, we immediately get the result by the Yoneda Lemma since we can just observe that

$$\text{Nat}(\text{Hom}(A, -), \text{Hom}(B, -)) \cong \text{Hom}(B, A).$$

Therefore, each such natural transformation is created from some $\varphi \in \text{Hom}(B, A)$, which is what we'd expect, so the Yoneda lemma passes our sanity check.

---

We now introduce the following definition to ease our discussion.

**Definition 3.2.9.** Let $\mathcal{C}$ be a category. A functor of the form $F : \mathcal{C}^{\text{op}} \longrightarrow \mathbf{Set}$ is called a **presheaf**[3]. As a presheaf may be viewed as an element of the functor category $\text{Fun}(\mathcal{C}^{\text{op}}, \mathbf{Set})$, we can define such a category as the **category of presheaves over $\mathcal{C}$**.

A natural source of presheaves is one which we are already familiar with. Given any locally small category $\mathcal{C}$, we can take any object $A$ of $\mathcal{C}$ to produce the functor

$$\text{Hom}_{\mathcal{C}}(-, A) : \mathcal{C}^{\text{op}} \longrightarrow \mathbf{Set}.$$

This process itself induces a functor known as the *Yoneda embedding*.

**Definition 3.2.10.** Let $\mathcal{C}$ be a locally small category. The **Yoneda embedding** on $\mathcal{C}$ is the functor $\boldsymbol{y} : \mathcal{C} \longrightarrow \text{Fun}(\mathcal{C}^{\text{op}}, \mathbf{Set})$ where for each object $A$

$$\boldsymbol{y}(A) = \text{Hom}_{\mathcal{C}}(-, A) : \mathcal{C}^{\text{op}} \longrightarrow \mathbf{Set}.$$

The reason why this is called the Yoneda embedding is because of the functor's relationship with the Yoneda embedding, which should become clear in proving the following proposition.

**Proposition 3.2.11.** The Yoneda embedding $\boldsymbol{y} : \mathcal{C} \longrightarrow \text{Fun}(\mathcal{C}^{\text{op}}, \mathbf{Set})$ is a full and faithful functor.

The proof of this proposition is left as an exercise. However, the Yoneda embedding arises naturally in many calculations within category. It is used to prove the following important proposition.

**Proposition 3.2.12.** Every small category $\mathcal{C}$ is concrete.

[3] The name "presheaf" is due to the fact that this concept is a precursor to the concept of a *sheaf*, which is outside of our scope for the moment.

*Proof.* Recall that a *concrete category* $\mathcal{C}$ is one which has a faithful functor $F : \mathcal{C} \longrightarrow \mathbf{Set}$. To demonstrate this for small categories, first define the functor

$$C : \text{Fun}(\mathcal{C}^{\text{op}}, \mathbf{Set}) \longrightarrow \mathbf{Set}$$

where a presheaf $P : \mathcal{C}^{\text{op}} \longrightarrow \mathbf{Set}$ is mapped as

$$(P : \mathcal{C}^{\text{op}} \longrightarrow \mathbf{Set}) \mapsto \coprod_{A \in \text{Ob}(\mathcal{C})} P(A).$$

Note that the indexing of the disjoint union is where we use locally smallness. This functor is fully faithful (exercise). As it is fully faithful, and the Yoneda embedding $\boldsymbol{y} : \mathcal{C} \longrightarrow \text{Fun}(\mathcal{C}^{\text{op}}, \mathbf{Set})$ is faithful, the composite functor

$$C \circ \boldsymbol{y} : \mathcal{C} \longrightarrow \mathbf{Set}$$

must be faithful. Hence we see that $\mathcal{C}$ is concrete. ∎

Finally, we end this section with a curious connection to group theory. It turns out that Yoneda's Lemma can actually be used in the proof of Cayley's Theorem. Sometimes this statement is taken too literally by others and they think "Yoneda's Lemma is a *generalization* of Cayley's Theorem" but that is simply not true, so the reader is warned to not believe someone when they hear that. Put simply, Yoneda's Lemma offers a bijection on sets which, with a little extra *separate* work, extends to an isomorphism of groups.

**Proposition 3.2.13.** (Cayley's Theorem.) Let $(G, \cdot)$ be a group. Then $G$ is isomorphic to a subgroup of $\text{Perm}(G)$.

*Proof.* Recall that a group $(G, \cdot)$ can be regarded as a category $\mathcal{C}$; specifically, we construct a category with one object $\bullet$ and set $\text{Hom}_{\mathcal{C}}(\bullet, \bullet) = U(G)$, where $U : \mathbf{Grp} \longrightarrow \mathbf{Set}$ is the forgetful functor. For each $g \in G$, a morphism is represented as $f_g : \bullet \longrightarrow \bullet$, and we have that $f_g \circ f_{g'} = f_{g' \cdot g}$.

Now consider the functor $\text{Hom}_{\mathcal{C}}(\bullet, -) : \mathcal{C} \longrightarrow \mathbf{Set}$. Such a functor produces the following data:

- We have that $\text{Hom}_{\mathcal{C}}(\bullet, \bullet) = U(G)$
- We also get a family of *bijections* $\varphi_g : U(G) \longrightarrow U(G)$ such that $\varphi_g \circ \varphi_{g'} = \varphi_{g' \cdot g}$.

In other words, the functor imposes an action of $G$ on its underlying set of elements $U(G)$ in $\mathbf{Set}$. Specifically, we may write $\varphi_{g'}(g) = g' \cdot g$ for each $g \in G$. Now what's a natural transformation $\eta$ between two functors?

$$\eta : \text{Hom}_{\mathcal{C}}(\bullet, -) \longrightarrow \text{Hom}_{\mathcal{C}}(\bullet, -).$$

Since there is only one object of $\mathcal{C}$, a natural transformation is *one* function $\eta : U(G) \longrightarrow U(G)$ such that for each $g' \in G$, the diagram below commutes.

$$
\begin{tikzcd}
\bullet \arrow[d, "f_{g'}"'] & U(G) \arrow[r, "\eta"] \arrow[d, "\varphi_{g'}"'] & U(G) \arrow[d, "\varphi_{g'}"] & g \arrow[r, mapsto] \arrow[d, mapsto] & \eta(g) \arrow[d, mapsto] \\
\bullet & U(G) \arrow[r, "\eta"'] & U(G) & g' \cdot g \arrow[r, mapsto] & \eta(g' \cdot g) = g' \cdot \eta(g)
\end{tikzcd}
$$

Now, Yoneda's Lemma gives us the bijection below, which we may denote as $\psi$,

$$\text{Nat}(\text{Hom}_{\mathcal{C}}(\bullet, -), \text{Hom}_{\mathcal{C}}(\bullet, -)) \cong \text{Hom}_{\mathcal{C}}(\bullet, \bullet) = U(G).$$

If we now observe that

- The collection of such natural transformations is a group under composition, with identity $1_{U(G)} : U(G) \longrightarrow U(G)$, which we may denote as $(P, \circ)$
- $(P, \circ) \subseteq \text{Perm}(G)$

then we can extend the isomorphism $\psi : P \longrightarrow U(G)$ to a group isomorphism

$$\psi : (P, \circ) \xrightarrow{\sim} (G, \cdot)$$

which is the statement of Cayley's Theorem. ∎

## Exercises

The first two exercises are very important. We (in fact you! The reader!) will use these results later on.

**1.** Prove the following dual counterpart to Proposition 3.2.1: Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor. Then a pair $(R, u : F(R) \longrightarrow D)$ is universal from $F$ to $D$ if and only if for each $C \in \mathcal{C}$, we have the natural bijection

$$\text{Hom}_{\mathcal{C}}(C, R) \cong \text{Hom}_{\mathcal{D}}(F(C), D).$$

**2.** Prove the following dual counterpart to Proposition 3.2.6: Let $\mathcal{C}$ be a locally small category, and suppose $K : \mathcal{C} \longrightarrow \textbf{Set}$ is a functor. Then $K$ is corepresentable, with representing object $R$, if and only if $(R, u : K(R) \longrightarrow \{\bullet\})$ is universal from $K$ to $\{\bullet\}$ for some object $R$.

*Hint:* Because $K$ is corepresentable, it is a contravariant functor. Thus, this should be very similar to the proof of Proposition 3.2.6, except with one twist.

# 3.3 Finite Products

In this section we will discuss products in categories, which will be our first encounter with the concept of a *limit*, something which has yet to be defined. The concept of a limit, and the dual concept of a colimit, form one of the central concepts of category theory. It will turn out that both the limit and colimit concepts are a special case of a universal morphism.

---

**Example 3.3.1.** Let $(G, \bullet)$ and $(H, \bullet)$ be two groups with group operations $\bullet : G \times G \longrightarrow G$ and $\bullet : H \times H \longrightarrow H$. The **product group** of $G, H$ is the group

$$(G \times H, \bullet) = \left\{ (g, h) \;\middle|\; g \in G,\ h \in H \right\}$$

whose group product works as

$$(g, h) \bullet (g', h') = (g \bullet g', h \bullet h').$$

One may check that this construction satisfies the definition of a group.

If $G, H$ are abelian groups, then the term "group product" is replaced with the term **direct sum** (we will explain why later). In this case, the product is denoted $(G \oplus H, \bullet)$, and the group operation does not change from above.

Direct sums, or more generally products of groups, are frequently used in group theory. For example, they are necessary to describe the fundamental theorem of finite abelian groups, which states that for any finite abelian group $A$, there exist primes $p_1, p_2, \dots, p_n$ and positive integers $\alpha_1, \alpha_2, \dots, \alpha_n$ such that

$$A \cong \mathbb{Z}_{p_1^{\alpha_1}} \oplus \mathbb{Z}_{p_2^{\alpha_2}} \oplus \cdots \oplus \mathbb{Z}_{p_n^{\alpha_n}}.$$

That is, every finite abelian group is the product of cycic groups of a prime-power order.

---

---

**Example 3.3.2.** Let $(X, \tau_X)$ and $(Y, \tau_Y)$ be two topological spaces. Using $X$ and $Y$, we can create a topological space $(X \times Y, \tau_{X \times Y})$ where $\tau_{X \times Y}$ is the **product topology**. There are many ways of defining this topology, but in the finite case, we can write $\tau_{X \times Y}$ as

$$\tau_{X \times Y} = \left\{ U \times V \;\middle|\; U \in \tau_X, V \in \tau_Y \right\}.$$

In the way we have presented this, this is actually the **box topology**, but the reader may recall that they coincide when we take finite products.

---

---

**Example 3.3.3.** In **Set**, we can always take two sets $X, Y$ to create the **cartesian product** $X \times Y$ defined as the set

$$X \times Y = \left\{ (x, y) \;\middle|\; x \in X, y \in Y \right\}$$

Now consider the following question.

> **Q:** What is the bare minimum amount of logical data that perfectly characterizes the above product $X \times Y$?

Well, observe that for such a set, we have *two* **projection functions**

$$p_1 : X \times Y \longrightarrow X \qquad p_1(x, y) = x$$
$$p_2 : X \times Y \longrightarrow Y \qquad p_2(x, y) = y.$$

Further, suppose that $f : Z \longrightarrow X$ and $g : Z \longrightarrow Y$ are two functions. Then there exists a third $h : Z \longrightarrow X \times Y$ such that $p_1 \circ h = f$ and $p_2 \circ h = g$. By this description, we can deduce that $h(z) = (f(z), g(z))$.

$$\begin{tikzcd} & Z \arrow[dl, "f"'] \arrow[d, dashed, "h"] \arrow[dr, "g"] & \\ X & X \times Y \arrow[l, "p_1"] \arrow[r, "p_2"'] & Y \end{tikzcd} \qquad \begin{tikzcd} & z \arrow[dl, mapsto] \arrow[d, mapsto] \arrow[dr, mapsto] & \\ f(z) & (f(z), g(z)) \arrow[l, mapsto] \arrow[r, mapsto] & g(z) \end{tikzcd} \tag{3.1}$$

Moreover, this $h$ is **unique** with respect to $f$ and $g$; Showing this is the bulk of Exercise 3.1.4. We now have an answer to our question.

> **A:** The product $X \times Y$ is characterized by the following data: two projection functions $p_1 : X \times Y \longrightarrow X, p_2 : X \times Y \longrightarrow Y$, such that for any pair of functions $f : Z \longrightarrow X, g : Z \longrightarrow Y$, there exists a **unique** third $h : Z \longrightarrow X \times Y$ such that diagram 3.1 commutes.

---

With the above example in mind, we now introduce our first definition of a product.

**Definition 3.3.4** (Nice Product Definition.)**.** Let $\mathcal{C}$ be a category with objects $A$ and $B$. The **product of** $A$ **and** $B$ is an object $A \times B$ equipped with morphisms

$$\pi_A : A \times B \longrightarrow A \qquad \pi_B : A \times B \longrightarrow B$$

with the following universal property: For any object $Z$ of $\mathcal{C}$ with morphisms $f : Z \longrightarrow A$, $g : Z \longrightarrow B$, there exists a unique morphism $h : Z \longrightarrow A \times B$ such that the diagram below commutes.

$$
\begin{tikzcd}
& Z \arrow[dl, "f"'] \arrow[d, dashed, "h"] \arrow[dr, "g"] & \\
A & A \times B \arrow[l, "\pi_A"] \arrow[r, "\pi_B"'] & B
\end{tikzcd}
$$

**Remark 3.3.5.** Note that to utilize the above universal property, one requires a *pair* of morphisms $f : Z \longrightarrow A$ and $g : Z \longrightarrow B$. That is, it is not true that, if I have a single morphism $f : Z \longrightarrow A$, then there exists a unique $h : Z \longrightarrow A \times B$ such that $\pi_A \circ h = k$. That would be false in many cases.

The above definition is a very nice one. For example, it returns the concepts of products of groups or topological spaces when it is imposed in **Grp** and **Top**. However, keep in mind the products don't always exist. For example, it does not work in **Fld**, the category of Fields (that is, there is no field which satisfies the universal property). We will eventually explain why.

---

**Example 3.3.6.** Consider **Ring**, the category of rings. We can create products in this category as follows: Let $(R, +, \bullet)$ and $(S, +, \bullet)$ be two rings with zeros $0_R, 0_S$ and units $1_R, 1_S$. Then we may form the **product ring** of $R$ and $S$ to be the ring

$$(R \times S, +, \bullet) = \left\{ (r, s) \;\middle|\; r \in R, s \in S \right\}$$

where for all pairs $(r_1, s_1)$ and $(r_2, s_2)$ in $R \times S$, we define the ring operations to behave as

- $(r_1, s_1) + (r_2, s_2) = (r_1 + r_2, s_1 + s_2)$
- $(r_1, s_1) \bullet (r_2, s_2) = (r_1 \bullet r_2, s_1 \bullet s_2)$

Note that with these requirements, the additive identity is $(0_R, 0_S)$ while the multiplicative identity is $(1_R, 1_S)$. With this construction, one can show that this satisfies the universal property of a product in **Ring**, so that **Ring** has products.

---

We make an interesting observation from the last example. For our ring $(R \times S, +, \bullet)$, we surely have that $(0_R, 1_S)$ and $(1_R, 0_R)$ are elements of the product ring. However,

$$(0_R, 1_S) \bullet (1_R, 0_S) = (0_R \bullet 1_R, 0_S \bullet 1_S) = (0_R, 0_S).$$

Hence, even if the rings $R$ and $S$ are integral domains, $R \times S$ is not an integral domain. Thus the product of two rings is never an integral domain.

---

**Example 3.3.7.** Consider the category of fields **Fld**. Let $F_1, F_2$ be fields. Then we would expect that the ring

$$F_1 \times F_2 = \left\{ (a, b) \;\middle|\; a \in F_1, F_2 \right\}$$

to be the "product field." But we just observed that this cannot be a field because the product ring is not even an integral domain.

However, this does not exclude the possibility that there is some kind of other field construction which we are not considering that plays the role as a product in **Fld**. We show that such a construction cannot hold for all fields with the following simple example.

Consider the fields $\mathbb{F}_2$ and $\mathbb{F}_3$, the fields with 2 and 3 elements, respectively. Suppose that $P$ is the product field of $\mathbb{F}_2$ and $\mathbb{F}_3$. Then by definition, we would require two projection field homomorphisms

$$\pi_1 : P \longrightarrow \mathbb{F}_2 \qquad \pi_2 : P \longrightarrow \mathbb{F}_3$$

However, recall that two fields share a (nonzero) field homomorphism if and only if they are of the same characteristic. Therefore,

- $\pi_1$ can only exist if $P$ has characteristic 2. In fact, $P$ must be isomorphic to $\mathbb{F}_2$.
- $\pi_2$ can only exist if it has characteristic 3. In fact, $P$ must be isomorphic $\mathbb{F}_3$.

Clearly, we have a contradiction. Thus we simply cannot generally take products in **Fld** in a logical way.

---

From the previous example, we see that products don't always exist in category. However, if they do, then we can take finitely many products. For instance, if we have three objects $A, B, C$, then we can take the products

$$A \times (B \times C) \qquad (A \times B) \times C.$$

If we have four objects, then we can create 5 products. Thus, if we can take the product of two objects, then we all finite products consisting of objects of $\mathcal{C}$ exist in our category.

We encapsulate this idea and include other prerequisites for a category to have finite products in the following proposition.

**Proposition 3.3.8.** Suppose $\mathcal{C}$ is a category with a terminal object $T$ and a product object $A \times B$ for every pair of objects $A$ and $B$. Then

***(i)*** $\mathcal{C}$ has finite products.

***(ii)*** There exists a bifunctor $\prod : \mathcal{C} \times \mathcal{C} \longrightarrow \mathcal{C}$ where $(A, B) \mapsto A \times B$.

***(iii)*** For any three objects, we have an isomorphism

$$(A \times B) \times C \cong A \times (B \times C)$$

which is natural in $A, B$ and $C$ .

***iv*** For any object $A$, we have the isomorphism

$$T \times A \cong A \cong T \times A$$

natural in $A$.

*Proof.* To prove the first part, let $P(n)$ be the following statement:

$$P(n) = \begin{cases} \text{For any objects } A_1, A_2, \dots, A_n \in \mathcal{C}, \\ \text{their product diagram in } \mathcal{C}. \end{cases}$$

**Base Case.** Observe that for $n = 0$, the statement is automatically true since we are given that a terminal object $T$ exists.

**Inductive Step.** Suppose the statement holds for $n = k$. Then for any objects $A_1, A_2, \cdots, A_k$, we have the product diagram

$$\begin{tikzcd}
 & & D \arrow[dll, bend right, "f_1"'] \arrow[dl, bend right, "f_2"'] \arrow[d, dashed, "u"] \arrow[dr, bend left, "f_{k-1}"] \arrow[drr, bend left, "f_k"] & & \\
 & & A_1 \times \cdots \times A_n \arrow[dll, "\pi_1"'] \arrow[dl, "\pi_2"'] \arrow[dr, "\pi_{k-1}"] \arrow[drr, "\pi_k"] & & \\
A_1 & A_2 & \cdots & A_{k-1} & A_k
\end{tikzcd}$$

and a unique, induced arrow $u$ whenever such a $D \in \mathcal{C}$ with morphisms $f_i : D \longrightarrow A_i$ exists.

Let $A_{k+1}$ be an arbitrary object of $\mathcal{C}$. Then the product $(A_1 \times A_2 \times \cdots \times A_k) \times A_{k+1}$ exists, since by assumption, the product of any two objects in our category must exist, and gives rise to the product diagram:

$$\begin{tikzcd}
 & D \arrow[dl, "g_1"'] \arrow[d, dashed, "v"] \arrow[dr, "g_2"] & \\
A_1 \times \cdots \times A_k & A_1 \times \cdots \times A_k \times A_{k+1} \arrow[l, "\pi'_1"] \arrow[r, "\pi'_2"'] & A_{k+1}
\end{tikzcd}$$

whenever such an object $D$ with a family of morphisms $g_1 : D \longrightarrow A_1 \times A_k$ and $g_2 : D \longrightarrow A_{k+1}$ exist.

Look at the bottom of the second diagram; we have a unique morphism $\pi'_1 : A_1 \times \cdots \times A_k \times A_{k+1} \longrightarrow A_1 \times \cdots \times A_k$. We can extend this across the morphisms $\pi_1, \pi_2 \cdots, \pi_k$ to demonstrate that there exist unique morphisms

$$\pi_i \circ \pi'_1 : A_1 \times \cdots \times A_k \times A_{k+1} \longrightarrow A_i$$

for $i = 1, 2, \dots, k$. Denote these as $\overline{\pi}_i$.

Now suppose we there exists an object $C$ in $\mathcal{C}$ with a family of morphisms $h_i : C \longrightarrow A_i$. Then by the first diagram, there exists a unique morphism $u : C \longrightarrow A_1 \times \cdots \times A_k$ such that $h_i = \pi_i \circ u$. Thus we have the diagram:

$$\begin{tikzcd}
 & C \arrow[dl, "u"'] \arrow[d, dashed, "v"] \arrow[dr, "h_{k+1}"] & \\
A_1 \times \cdots \times A_k & A_1 \times \cdots \times A_k \times A_{k+1} \arrow[l, "\pi_1'"] \arrow[r, "\pi_2'"'] & A_{k+1}
\end{tikzcd}$$

so we have a unique morphism $v : C \longrightarrow A_1 \times \cdots \times A_{k+1}$ such that $\pi_1' \circ v = u$ and $\pi_2' \circ v = h_{k+1}$. However, note that

$$\pi_1' \circ v = u \implies (\pi_i \circ \pi_1') \circ v = \pi_i \circ u \implies \overline{\pi}_i \circ v = h_i.$$

for $i = 1, 2, \dots, k$.

Now let $\overline{\pi}_{k+1} = \pi_2'$. Then we see that for such a family $h_i : C \longrightarrow A_i$ for $i = 1, 2, \dots, k+1$, there exists a unique morphism $v : C \longrightarrow A_1 \times \cdots \times A_{k+1}$ such that

$$\overline{\pi}_i \circ v = h_i$$

for $i = 1, 2, \dots, k+1$. Therefore, we have the product diagram

$$\begin{tikzcd}
 & & C \arrow[dll, bend right, "h_1"'] \arrow[dl, bend right, "h_2"'] \arrow[d, dashed, "v"] \arrow[dr, bend left, "h_k"] \arrow[drr, bend left, "h_{k+1}"] & & \\
 & & A_1 \times \cdots \times A_{k+1} \arrow[dll, "\overline{\pi}_1"] \arrow[dl, "\overline{\pi}_2"] \arrow[dr, "\overline{\pi}_k"'] \arrow[drr, "\overline{\pi}_{k+1}"] & & \\
A_1 & A_2 & \cdots & A_k & A_{k+1}
\end{tikzcd}$$

so that the product $A_1 \times A_k \times A_{k+1}$ exists and is well-defined in $\mathcal{C}$. Hence, $P(n)$ is true for $n = k+1$.

By mathematical induction, we see that all finite products must exist in $\mathcal{C}$, as desired.

To demonstrate the existence of a bifunctor, we can directly define one. Let $\prod : \mathcal{C} \times \mathcal{C} \longrightarrow \mathcal{C}$ act as follows.

**Objects.** $\prod(A, B) = A \times B$.

**Morphisms.** Let $f : A \longrightarrow A'$ and $g : B \longrightarrow B'$. Suppose we have canonical projections

$$\pi_1 : A \times B \longrightarrow A \qquad \pi_2 : A \times B \longrightarrow B$$

and

$$\pi_1' : A' \times B' \longrightarrow A' \qquad \pi_2' : A' \times B' \longrightarrow B'.$$

Then observe we get the diagram

$$\begin{tikzcd}
 & A \times B \arrow[dl, "f \circ \pi_1"'] \arrow[d, dashed, "u"] \arrow[dr, "g \circ \pi_2"] & \\
A' & A' \times B' \arrow[l, "\pi_1'"] \arrow[r, "\pi_2'"'] & B'
\end{tikzcd}$$

Thus, there exists a unique morphism $u : A \times B \longrightarrow A' \times B'$ whenever such $f, g$ exist. Therefore, we can define how $\prod$ acts on morphism as

$$\prod(f : A \longrightarrow A', g : B \longrightarrow B') = u : A \times B \longrightarrow A' \times B'$$

where $u$ is generated by the diagram above. As we just showed, this assignment is well-defined. It's now pretty straightforward to now show that this establishes a functor (and I'm too lazy to do so).
To establish associativity of our products, we demonstrate they're isomorphic. Thus let $A \times (B \times C)$ and $(A \times B) \times C$ be two products in $\mathcal{C}$. Suppose we have an family of morphisms $h_1 : D \longrightarrow A$, $h_2 : D \longrightarrow B$ and $h_3 : D \longrightarrow C$. Then we get the following product diagrams.

$$\begin{tikzcd}
 & D \arrow[dl, "h_2"'] \arrow[d, dashed, "v"] \arrow[dr, "h_3"] & \\
B & B \times C \arrow[l, "p_1"] \arrow[r, "p_2"'] & C
\end{tikzcd}$$

$$\begin{tikzcd}
 & D \arrow[dl, "h_1"'] \arrow[d, dashed, "w"] \arrow[dr, "h_2"] & \\
A & A \times B \arrow[l, "p_1'"] \arrow[r, "p_2'"'] & B
\end{tikzcd}$$

$$\begin{tikzcd}
 & D \arrow[d, dashed, "u"] \arrow[ddl, bend right, "h_1"'] \arrow[dd, bend right, "h_2"] \arrow[ddr, bend left, "h_3"] & \\
 & A \times B \times C \arrow[dl, "\overline{\pi}_1"'] \arrow[d, "\overline{\pi}_2"] \arrow[dr, "\overline{\pi}_3"] & \\
A & B & C
\end{tikzcd}$$

Since we have unique morphisms $v : D \longrightarrow B \times C$ and $w : D \longrightarrow A \times B$, we also get the product diagrams.

$$\begin{tikzcd}
 & D \arrow[dl, "h_1"'] \arrow[d, dashed, "y"] \arrow[dr, "v"] & \\
A & A \times (B \times C) \arrow[l, "\pi_1"] \arrow[r, "\pi_2"'] & B \times C
\end{tikzcd}
\qquad
\begin{tikzcd}
 & D \arrow[dl, "w"'] \arrow[d, dashed, "z"] \arrow[dr, "h_3"] & \\
A \times B & (A \times B) \times C \arrow[l, "\pi_1'"] \arrow[r, "\pi_2'"'] & C
\end{tikzcd}$$

for the products $A \times (B \times C)$ and $(A \times B) \times C$, respectively. Thus we have the collection of morphisms

$$\begin{aligned}
&p_1' \circ \pi_1' : (A \times B) \times C \longrightarrow A && \pi_1 : A \times (B \times C) \longrightarrow A \\
&p_2' \circ \pi_1' : (A \times B) \times C \longrightarrow B && p_1 \circ \pi_2 : A \times (B \times C) \longrightarrow B \\
&\pi_2' : (A \times B) \times C \longrightarrow C && p_2 \circ \pi_2 : A \times (B \times C) \longrightarrow C.
\end{aligned}$$

Now observe that

$$p_1' \circ \pi_1' \circ z = p_1' \circ w = h_1 \qquad \pi_1 \circ y = h_1 \tag{3.2}$$

$$p_2' \circ \pi_1' \circ x = p_2' \circ w = h_2 \qquad p_1 \circ \pi_2 \circ y = p_1 \circ v = h_2 \tag{3.3}$$

$$\pi_2' \circ z = h_3 \qquad p_2 \circ \pi_2 \circ y = p_2 \circ v = h_3. \tag{3.4}$$

Thus we see that our first collection of morphisms are projections. That is, for any family of morphisms $h_1 : D \longrightarrow A$, $h_2 : D \longrightarrow B$ and $h_3 : D \longrightarrow C$, there exists unique morphisms such that equations $y, z$ such that equations (3), (4) and (5) hold. What this means is that $A \times (B \times C)$ and $(A \times B) \times C$ are universal objects; specifically, they form universal cones. However, the original universal cone of this construction was simply $A \times B \times C$ with the morphisms $\overline{\pi}_1, \overline{\pi}_2, \overline{\pi}_3$. Thus we have that

$$A \times (B \times C) \cong (A \times B) \times C \cong A \times B \times C$$

since universal objects of the same construction are isomorphic. Showing naturality is not hard (again, too lazy to do that).

Finally, let $T$ be the terminal object of $\mathcal{C}$. Denote $t_C : C \longrightarrow T$ as the unique morphism from $C$ to $T$. Now consider the product diagram associated with the product $T \times A$:

$$\begin{tikzcd} & D \arrow[dl, "t_D"'] \arrow[d, dashed, "u"] \arrow[dr, "f"] & \\ T & T \times A \arrow[l, "t_A"] \arrow[r, "\pi"'] & A \end{tikzcd}$$

Observe that $t_D$ always exists for any $D$. Hence the existence of $u$ is completely dependent $f$. Therefore, we can see that this diagram is equivalent to

$$\begin{tikzcd} & D \arrow[dl, "t_D"'] \arrow[d, dashed, "u"] \arrow[dr, "f"] & \\ T & A \arrow[l, "t_A"] \arrow[r, "1_A"'] & A \end{tikzcd}$$

Hence we see that $A$ with the morphism $t_A, 1_A$ forms a universal cone. But so does $T \times A$; hence, uniqueness guarantees they are isomorphic.

∎

Now that we have discussed examples of products in categories, offered a rigorous definition, and we observed an example when they do not exist, we would naturally want to generalize this concept since it is often the case that we would like to take arbitrary products, or even infinite products. We also want to somehow connect products to a universal morphism. To do all of these things requires us to further abstract our definition of a product. Before doing so, we offer a simple definition.

**Definition 3.3.9.** Let $\mathcal{C}$ be a category. Define the **diagonal functor of $\mathcal{C}$** as $\Delta : \mathcal{C} \longrightarrow \mathcal{C} \times \mathcal{C}$ where

**On Objects.** For $C$ an object of $\mathcal{C}$, we define $\Delta(C) = (C, C)$.

**On Morphisms.** For a morphism $f : A \longrightarrow B$, we define $\Delta(f) = (f, f) : (A, A) \longrightarrow (B, B)$.

The above functor is a bit silly; it really doesn't do much. However, it necessary for us to really understand what exactly a product is. It helps us realize that a product in a category is actually a universal morphism.

**Definition 3.3.10** (Rigorous Product Definition.)**.** Let $\mathcal{C}$ be a category with objects $A, B$. The **product** $A \times B$ **of** $A$ **and** $B$ is a universal morphism

$$\pi : (A \times B, u : \Delta(A \times B) \longrightarrow (A, B))$$

from $\Delta$ to $(A, B)$. This means that for any other pair $(C, q : \Delta(C) \longrightarrow (A, B))$, there exists a unique $h : C \longrightarrow A \times B$ in $\mathcal{C}$ such that the diagram below commutes.

$$\begin{tikzcd}
\Delta(A \times B) \arrow[r, "\pi"] & (A, B) \\
\Delta(C) \arrow[u, dashed, "\Delta(h)"] \arrow[ur, "q"'] &
\end{tikzcd}$$

This definition is exactly equivalent to our previous. What this tells us is that a product is an instance of a universal morphism. We show how this definition is equivalent to the previous via the following example.

---

**Example 3.3.11.** To see this for the case when $n = 2$, consider the product $A \times B$ of two objects $A, B$ in some category $\mathcal{C}$. Then

$$\begin{tikzcd}
(A, B) & \Delta(A \times B) \arrow[l, "u"'] \\
& \Delta(C) \arrow[ul, "h"] \arrow[u, dashed, "\Delta(f')"']
\end{tikzcd}
=
\begin{tikzcd}
(A, B) & (A \times B, A \times B) \arrow[l, "(\pi_A, \pi_B)"'] \\
& (C, C) \arrow[ul, "h"] \arrow[u, dashed, "(f', f')"']
\end{tikzcd}
\qquad
\begin{tikzcd}
A \times B \\
C \arrow[u, dashed, "f'"]
\end{tikzcd}$$

Let's spell out what's going on above; you might have seen this exposition, without even realizing, demonstrating the universality of products. Suppose there exists another object $C$ with morphisms $f : C \longrightarrow A$ and $g : C \longrightarrow B$. Then we force the existence of a morphism $f' : C \longrightarrow A \times B$.

$$\begin{tikzcd}
A & A \times B \arrow[l, "\pi_A"'] \arrow[r, "\pi_B"] & B \\
& C \arrow[ul, "f"] \arrow[u, dashed, "f'"] \arrow[ur, "g"'] &
\end{tikzcd}$$

When we usually do this, we simply just set

$$f' = (f, g)$$

so that $\pi_A \circ f' = f$, and $\pi_B \circ f' = g$.

---

## 3.4 Finite Coproducts

We now move onto the concept of coproducts in categories. We will see that this concept is an instance of a *colimit*, which is yet to be defined. We build intuition on the concept with the special concept of coproducts by introducing examples.

---

**Example 3.4.1.** Let $(G, \cdot)$ and $(H, \cdot)$ be two groups with group operations $\cdot : G \times G \longrightarrow G$ and $\cdot : H \times H \longrightarrow H$. The **free product** of $G$ and $H$ is the group

$$(G * H, \bullet) = \left\{ g_1 h_1 g_2 h_2 \cdots g_k h_k \;\middle|\; g_i \in G, h_i \in H \right\}$$

with the following operation. If $g_1 h_1 \cdots g_k h_k$ and $g_1' h_1' \cdots g_\ell' h_\ell'$ are two elements of $G * H$, then

$$(g_1 h_1 \cdots g_k h_k) \bullet (g_1' h_1' \cdots g_\ell' h_\ell') = g_1 h_1 \cdots g_k h_k g_1' h_1' \cdots g_\ell' h_\ell'.$$

We require the group operation to obey the following two rules. Let $g_1 h_1 \cdots g_k h_k \in G * H$.

- If $g \in G$, then
$$g \bullet (g_1 h_1 \cdots g_k h_k) = (g \cdot g_1) h_1 \cdots g_k h_k.$$

- If $h \in H$, then
$$(g_1 h_1 \cdots g_k h_k) \bullet h = g_1 h_1 \cdots g_k (h_k \cdot h).$$

The free product of two groups arise frequently in algebraic topology. Despite that its definition is somewhat complicated, we will see later that free products are in some sense dual to the concept of the product of groups. The reader will also soon see that the naming "free product" is an unfortunate one as it is somewhat misleading.

Free products appear prominently in various statements of Van Kampen's theorem in topology; what follows is a simplified version. If $X = U \cup V$ is a topological space with $U, V$ open sets, and if $U \cap V \neq \varnothing$ is path connected and simply connected, then

$$\pi_1(X) \cong \pi_1(U) * \pi_1(V)$$

where $\pi_1(X)$ is the fundamental group of $X$. (Note that since $X$ is path connected, it doesn't matter what basepoint for the fundamental group we select).

---

We will soon see that the free product is the coproduct in the category of **Grp**, although such a statement should not make any sense the reader until we define what a coproduct is.

---

**Example 3.4.2.** In **Set**, we can combine two different sets $X$ and $Y$ to create the **disjoint**

**union** $X \amalg Y$, which is defined to be the set

$$X \amalg Y = \left\{ (x, 0), (y, 1) \;\middle|\; x \in X, y \in Y \right\}.$$

In the above set, elements are tuples whos first coordinate is either in $X$ or $Y$, and the second is some value which depends on whether or not the first coordinate is in $X$ or $Y$. I decided to make these values 0 and 1, but it is ultimately arbitrary. We just need to make sure that these values are distinct so that we can determine if a tuple has an element from $X$ or $Y$ based on the value in the second slot. For example, for a tuple $(z, 0)$, we know that $z \in X$. If the tuple is of the form $(z, 1)$, we know that $z \in Y$.

We perform a similar analysis as before with products, and we consider the following question.

> **Q:** What is the bare minimum amount of logical data that perfectly characterizes the above disjoint union $X \amalg Y$?

Observe that we have the two **inclusion functions**

$$\begin{aligned} i_1 &: X \longrightarrow X \amalg Y \qquad i_1(x) = (x, 0) \\ i_2 &: Y \longrightarrow X \amalg Y \qquad i_2(y) = (y, 1). \end{aligned}$$

These two functions are equipped with the following remarkable property. Let $Z$ be some set, and suppose I have two functions

$$\begin{aligned} f &: X \longrightarrow X \amalg Y \\ g &: Y \longrightarrow X \amalg Y. \end{aligned}$$

Then there exists a unique function $h : X \amalg Y \longrightarrow Z$ such that the diagram below commutes.

$$\begin{tikzcd} & Z & \\ X \arrow[ur, "f"] \arrow[r, "i_1"'] & X \amalg Y \arrow[u, dashed, "h"] & Y \arrow[ul, "g"'] \arrow[l, "i_2"] \end{tikzcd} \qquad h(z, i) = \begin{cases} f(z) & \text{if } i = 0 \\ g(z) & \text{if } i = 1 \end{cases} \tag{3.5}$$

This definition of this unique $h : X \amalg Y \longrightarrow Z$ is described above on the right. With the above definition, one can easily see that the above diagram does in fact commute. We now have an answer to our question.

> **A:** The disjoint union $X \amalg Y$ is characterized by two inclusion functions $i_1 : X \longrightarrow X \amalg Y$, $i_2 : Y \longrightarrow X \amalg Y$, such that, for any $f : X \longrightarrow Z$, $g : Y \longrightarrow Z$, there exists a **unique** $h : X \amalg Y \longrightarrow Z$ such that diagram 3.5 commutes.

---

This now motivates the following definition of a *coproduct*.

**Definition 3.4.3** (Nice Coproduct Definition.)**.** Let $\mathcal{C}$ be a category with objects $A$ and $B$. The **coproduct** of $A$ and $B$ is an object $A \amalg B$ of $\mathcal{C}$ which is equipped with morphisms

$$i_A : A \longrightarrow A \amalg B \qquad i_B : B \longrightarrow A \amalg B$$

with the following universal property: For any object $Z$ of $\mathcal{C}$ with a pair of morphisms $f : A \longrightarrow Z$ and $g : B \longrightarrow Z$, then there exists a unique morphism $h : A \amalg B \longrightarrow Z$ such that the diagram below commutes.

$$\begin{tikzcd} & Z & \\ A \arrow[ur, "f"] \arrow[r, "i_A"'] & A \amalg B \arrow[u, dashed, "h"] & B \arrow[ul, "g"'] \arrow[l, "i_B"] \end{tikzcd}$$

It is now clear that, coproducts in **Set** exist; it is the disjoint union.

**Remark 3.4.4.** Note that to utilize the above universal property, one requires a *pair* of morphisms $f : A \longrightarrow Z$ and $g : B \longrightarrow Z$. That is, it is not true that, if I have a single morphism $k : A \longrightarrow Z$, then there exists a unique $h : A \amalg B \longrightarrow Z$ such that $h \circ i_X = k$. That would be false in many cases.

**Proposition 3.4.5.** Suppose $\mathcal{C}$ is a category with an initial object $I$ and a coproduct object $A \amalg B$ for every pair of objects $A$ and $B$. Then

***(i)*** $\mathcal{C}$ has finite coproducts.

***(ii)*** There exists a bifunctor $\amalg : \mathcal{C} \times \mathcal{C} \longrightarrow \mathcal{C}$ where $(A, B) \mapsto A \amalg B$.

***(iii)*** For any three objects, we have an isomorphism

$$(A \amalg B) \amalg C \cong A \amalg (B \amalg C) \cong A \amalg B \amalg C$$

which is natural in $A, B$ and $C$ .

***iv*** For any object $A$, we have the isomorphism

$$I \amalg A \cong A \cong I \amalg A$$

natural in $A$, where $T$ is the initial object of the category.

**Definition 3.4.6** (Rigorous Coproduct Definition)**.** Let $\mathcal{C}$ be a category with objects $A, B$. The **coproduct** $A \amalg B$ **of** $A$ **and** $B$ is a universal morphism

$$(A \amalg B, i : (A, B) \longrightarrow \Delta(A \amalg B))$$

from $(A, B)$ to $\Delta$. This means that, for any other pair $(C, j : (A, B) \longrightarrow \Delta(C))$, there exists a unique $h : A \amalg B \longrightarrow C$ such that the diagram below commutes.

$$
\begin{tikzcd}
(A,B) \arrow[r, "i"] \arrow[dr, "j"'] & \Delta(A \amalg B) \arrow[d, dashed, "\Delta(h)"] & A \amalg B \arrow[d, dashed, "h"] \\
& \Delta(C) & C
\end{tikzcd}
$$

Visually, we have that

$$
\begin{tikzcd}
(A,B) \arrow[r, "u"] \arrow[dr, "f"'] & (C,C) \arrow[d, dashed, "(h{,}h)"] & C \arrow[d, dashed, "h"] \\
& (D,D) & D
\end{tikzcd}
$$

## 3.5 Arbitrary Products and Coproducts in Categories

In this section, we perform a construction that allows us to have finite products and coproducts in a category. Once we achieve that construction, we easily generalize our work to obtain a definition for arbitrary products and coproducts in a category.

**Definition 3.5.1.** Let $\mathcal{D}_n$ be the discrete category with $n$-many objects (we use the letter $\mathcal{D}$ for "discrete"). We will often visualize $\mathcal{D}_n$ as below.

$$\begin{tikzcd} \bullet_1 \arrow[loop above] & \bullet_2 \arrow[loop above] & \bullet_3 \arrow[loop above] & \cdots & \bullet_n \arrow[loop above] \end{tikzcd}$$

Note that a functor $F : \mathcal{D}_n \longrightarrow \mathcal{C}$ is one which simply picks out $n$ different objects $A_1$, $A_2, \dots, A_n$ of $\mathcal{C}$:

$$F(\bullet_1) = A_1, \quad F(\bullet_2) = A_2, \quad \dots, \quad F(\bullet_n) = A_n.$$

This category allows us to make the following definition.

**Definition 3.5.2.** Let $\mathcal{C}$ be a category. The $n$-th **diagonal functor** $\Delta_n : \mathcal{C} \longrightarrow \mathcal{C}^n$, is the functor defined as follows.

**On Objects.** For an object $C$, we have that $\Delta_n(C) = (\overbrace{C, C, \dots, C}^{n\text{-many copies}})$.
**On Morphisms.** For a morphism $f : A \longrightarrow B$ in $\mathcal{C}$, we have that

$$\Delta_n(f : A \longrightarrow B) = (f, f, \dots, f) : \Delta_n(A) \longrightarrow \Delta_n(B).$$

The diagonal functor is also sometimes informally called the "copy" functor, since it is literally just copying data. We now make some observations.

(**1**) For each object $C \in \mathcal{C}$, we can interpret the object $\Delta_n(C) \in \mathcal{C}^n$ *as a functor*

$$\Delta_n(C) : \mathcal{D}_n \longrightarrow \mathcal{C}$$

where $\Delta_n(C)$ sends $\bullet_i$ to $C$ for all $i = 1, 2, \dots, n$.

(**2**) Thus, we may also regard the $n$-th diagonal functor as a functor as below.

$$\Delta_n : \mathcal{C} \longrightarrow \mathrm{Fun}(\mathcal{D}_n, \mathcal{C}) \qquad C \mapsto (\Delta_n(C) : \mathcal{D}_n \longrightarrow \mathcal{C}).$$

In this interpretation, every morphism $f : C \longrightarrow C'$ is interpreted as a natural transformation $\Delta_n(f) : \Delta_n(C) \longrightarrow \Delta_n(C')$.

(**3**) Consider a functor $F : \mathcal{D}_n \longrightarrow \mathcal{C}$ such that $F(\bullet_i) = A_i \in \mathcal{C}$. For each $C \in \mathcal{C}$, a natural transformation

$$\eta : \Delta_n(C) \longrightarrow F$$

will simply correspond to $n$-many morphisms $\eta_1, \dots, \eta_n$ where

$$\eta_i : \Delta_n(C)(\bullet_i) \longrightarrow F(\bullet_i) \implies \eta_i : C \longrightarrow A_i.$$

With this notation clarified, we now can propose our definition of a product.

**Definition 3.5.3** (Finite Product and Coproduct Definition)**.** Let $\mathcal{C}$ be a category. Let $A_1$, $A_2, \dots,$ $A_n$ be objects of $\mathcal{C}$. Let $F : \mathcal{D}_n \longrightarrow \mathcal{C}$ be the functor such that $F(\bullet_i) = A_i$.

- The **product** of $A_1, A_2, \dots, A_n$ is an object $P$ of $\mathcal{C}$ equipped with a (natural transformation) $p : \Delta_n(P) \longrightarrow F$ such that

$$(P, p : \Delta_n(P) \longrightarrow F) \text{ is universal from } \Delta_n \text{ to } P.$$

In the case where the product $P$ exists, we write $P = \prod_{i=1}^{n} A_i$.

- The **coproduct** of $A_1, A_2, \dots, A_n$ is an object $C$ of $\mathcal{C}$ equipped with a (natural transformation) $i : F \longrightarrow \Delta_n(C)$ such that

$$(C, i : F \longrightarrow \Delta_n(C)) \text{ is universal from } C \text{ to } \Delta_n.$$

In the case where the coproduct $C$ exists, we write $C = \coprod_{i=1}^{n} A_i$

**Remark 3.5.4.** By Observation (**3**) as above, if $p : \Delta_n\left(\prod_{i=1}^{n} A_i\right) \longrightarrow F$ is a natural transformation, then it corresponds to $n$-many morphisms

$$p_k : \prod_{i=1}^{n} A_i \longrightarrow A_k \qquad k = 1, 2, \dots, n$$

This matches our intuition: A product of $n$-objects should always have $n$-many morphisms between the product and each of its factors.

Similarly, a natural transformation $i : F \longrightarrow \Delta_n\left(\coprod_{i=1}^{n} A_i\right)$ corresponds to $n$-many morphisms

$$i_k : A_k \longrightarrow \coprod_{i=1}^{n} A_i \qquad k = 1, 2, \dots, n$$

which again matches our intuition: A coproduct of $n$-objects should have $n$-many morphisms between each of its factors and the coproduct.

We now have everything we need to define arbitrary products and coproducts, including infinite ones. We just need to specify some notation that we will use. Towards that goal, let $\lambda$ be some indexing set.

- Define $\mathcal{D}_\lambda$ to be the discrete category consisting of one object $\bullet_i$ for each $i \in \lambda$. (In particular, $\mathcal{D}_\lambda$ is now possibly infinite.)

- Define the $\lambda$-diagonal functor to be the functor $\Delta_\lambda : \mathcal{C} \longrightarrow \text{Fun}(\mathcal{D}_\lambda, \mathcal{C})$ where $\Delta_\lambda(C) : \mathcal{D}_\lambda \longrightarrow \mathcal{C}$ sends each $\bullet_i$ to $C$ for all $i \in \lambda$.

**Definition 3.5.5** (Arbitrary Product and Coproduct Definition)**.** Let $\mathcal{C}$ be a category, and consider an arbitrary set of objects $\{A_i\}_{i\in\lambda}$ of $\mathcal{C}$, $\lambda$ some indexing set. Let $F : \mathcal{D}_\lambda \longrightarrow \mathcal{C}$ be the functor such that $F(\bullet_i) = A_i$ for $i \in \lambda$.

- The **product** of $\{A_i\}_{i\in\lambda}$ is the object $P$ of $\mathcal{C}$ equipped with a (natural transformation) $p : \Delta_\lambda(P) \longrightarrow F$ such that

$$(P, \Delta_\lambda(P) \longrightarrow F) \text{ is universal from } \Delta_\lambda \text{ to } P.$$

  In the case where the product $P$ exists, we write $P = \prod_{i\in\lambda} A_i$.

- The **coproduct** of $\{A_i\}_{i\in\lambda}$ is the object $C$ of $\mathcal{C}$ equipped with a (natural transformation) $i : F \longrightarrow \Delta_\lambda(C)$ such that

$$(C, i : F \longrightarrow \Delta_\lambda(C)) \text{ is universal from } C \text{ to } \Delta_\lambda.$$

**Remark 3.5.6.** Notice the inherent duality present in the definition of a product and coproduct. This is one of the reasons category theory is nice; one now has a new perspective of understanding, for example, the free product operation and the group product operation in **Grp**; they're dual concepts!

Since products and coproducts of objects are universal objects, we obtain some nice results since we already know how universal objects operate. Before introduce such results, we require the following lemma.

**Lemma 3.5.7.** Let $\mathcal{C}$ be a locally small category, and let $\{A_i\}_{i\in\lambda}$ be objects of $\mathcal{C}$. Suppose their product exists in $\mathcal{C}$. Then the functor

$$\prod_{i\in\lambda} \text{Hom}_{\mathcal{C}}(-, A_i) : \mathcal{C} \longrightarrow \textbf{Set}$$

which sends an object $C$ to the set $\prod_{i\in\lambda} \text{Hom}_{\mathcal{C}}(C, A_i)$ is representable by the functor

$$\text{Hom}_{\text{Fun}(\mathcal{D}_\lambda, \mathcal{C})}(\Delta_\lambda(-), F) : \mathcal{C} \longrightarrow \textbf{Set}.$$

The proof is left as an exercise. It is not difficult to show; it simply requires realizing that there is a natural bijection between $\prod_{i\in\lambda} \text{Hom}_{\mathcal{C}}(C, A_i)$ and $\text{Hom}_{\text{Fun}(\mathcal{D}_\lambda,\mathcal{C})}(\Delta_\lambda(C), F)$ for each $C \in \mathcal{C}$.

Using all of our previous work we now have the following proposition.

**Proposition 3.5.8.** Let $\mathcal{C}$ be a locally small category, and let $\{A_i\}_{i\in\lambda}$ be a set of objects in $\mathcal{C}$. Denote $F : \mathcal{D}_\lambda \longrightarrow \mathcal{C}$ where $F(\bullet_i) = A_i$ for all $i \in \lambda$.

- If the product $\prod_{i\in\lambda} A_i$ exists in $\mathcal{C}$, then for each object $C$ of $\mathcal{C}$, we have the natural bijection
$$\prod_{i\in\lambda} \text{Hom}_{\mathcal{C}}(C, A_i) \cong \text{Hom}_{\mathcal{C}}\left(C, \prod_{i\in\lambda} A_i\right)$$
- If the coproduct $\coprod_{i\in\lambda} A_i$ exists in $\mathcal{C}$, then for each object $C$ of $\mathcal{C}$, we have the natural bijection
$$\prod_{i\in\lambda} \text{Hom}(A_i, C) \cong \text{Hom}_{\mathcal{C}}\left(\coprod_{i\in\lambda} A_i, C\right).$$

*Proof.* We only prove the first result, since the second follows similarly. Since $\prod_{i\in\lambda} A_i$ exists in $\mathcal{C}$, we know that this implies $\prod_{i\in\lambda} A_i$ is equipped with a natural transformation $p : \Delta_\lambda\left(\prod_{i\in\lambda} A_i\right) \longrightarrow F$ such that $\left(\prod_{i\in\lambda} A_i, p\right)$ is universal from $\Delta_\lambda$ to $P$.

From this perspective, we can apply the result of Exercise 3.2.1 to conclude that, for each object $C$, we have the natural bijection below.
$$\text{Hom}_{\mathcal{C}}\left(C, \prod_{i\in\lambda} A_i\right) \cong \text{Hom}_{\text{Fun}(\mathcal{D}_\lambda, \mathcal{C})}(\Delta_\lambda(C), F).$$
However, we know from Lemma 3.5.7 that there is a natural bijection
$$\text{Hom}_{\text{Fun}(\mathcal{D}_\lambda, \mathcal{C})}(\Delta_\lambda(C), F) \cong \prod_{i\in\lambda} \text{Hom}_{\mathcal{C}}(C, A_i).$$
Thus we have a natural bijection
$$\prod_{i\in\lambda} \text{Hom}_{\mathcal{C}}(C, A_i) \cong \text{Hom}_{\mathcal{C}}\left(C, \prod_{i\in\lambda} A_i\right)$$
as desired.

The second result is left as an exercise (we outline the steps for the reader). ∎

**Remark 3.5.9.** Note that the above proposition is saying something very deep and beautiful about products and coproducts as a concept. Moreover, also note that a direct proof would have been very long-winded and complicated, but that our previous work made it possible to give a proof consisting of a few lines. Thus, a categorical perspective is evidently sometimes useful.

We now introduce the following interesting property. This property becomes an important observation when we begin look at *abelian categories*.

**Proposition 3.5.10.** Let $\mathcal{C}$ be a category and let $\{A_i\}_{i\in\lambda}$ be a set of objects in $\mathcal{C}$. Suppose the product $\prod_{i\in\lambda} A_i$ and coproduct $\coprod_{i\in\lambda} A_i$ exist in $\mathcal{C}$. Then there is a canonical morphism

$$\varphi : \prod_{i\in\lambda} A_i \longrightarrow \coprod_{i\in\lambda} A_i$$

in $\mathcal{C}$.

*Proof.* Let $F : \mathcal{D}_\lambda \longrightarrow \mathcal{C}$ be the functor where $F(\bullet_i) = A_i$. Then the product and coproduct are equipped with the natural transformations as below.

$$\Delta_\lambda\left(\prod_{i\in\lambda} A_i\right) \longrightarrow F \qquad F \longrightarrow \Delta_\lambda\left(\coprod_{i\in\lambda} A_i\right)$$

Then we can compose them to obtain the natural transformation

$$\Delta_\lambda\left(\prod_{i\in\lambda} A_i\right) \longrightarrow \Delta_\lambda\left(\coprod_{i\in\lambda} A_i\right).$$

By the universal property of the coproduct, this implies a unique $\varphi : \prod_{i\in\lambda} A_i \longrightarrow \coprod_{i\in\lambda} A_i$ such that the diagram below commutes.

∎

**Remark 3.5.11.** Here is one of our first uses of the word "canonical." This is not an adjective that adds detail to our morphism (e.g., an extra mathematical property), but it is a word we superfluously wrote to emphasize to the reader that morphisms of a given form cannot always be found in categories.

The word "canonical" is often used in category theory language, but it is never really defined because its always secretly assumed that everyone knows what it means. It's a useful word, so we will use it later on, but again: it means nothing more than "There exists an obvious morphism of a given form."

## Exercises

1. Prove Lemma 3.5.7. (Note: the notation and statement may make it look harder than it actually is.)
2. Complete the proof of Proposition 3.5.8 as follows.

   *i.* Show that the functor
   $$\prod_{i\in\lambda} \mathrm{Hom}_{\mathcal{C}}(A_i, -) : \mathcal{C} \longrightarrow \mathbf{Set}$$
   is representable by the functor
   $$\mathrm{Hom}_{\mathrm{Fun}(\mathcal{D}_\lambda,\mathcal{C})}(F, \Delta_\lambda(-)) : \mathcal{C} \longrightarrow \mathbf{Set}$$

*ii.* Using $(i)$, Proposition 3.2.1, and interpreting coproducts as universal objects, prove that

$$\prod_{i\in\lambda} \mathrm{Hom}(A_i, C) \cong \mathrm{Hom}_{\mathcal{C}}\left(\coprod_{i\in\lambda} A_i,\ C\right).$$

**3.** Let $P$ be a preorder with binary relation $\leq$. Consider a subset $A \subseteq P$ where $A = \{a_i \in P \mid i \in \lambda\}$ with $\lambda$ some indexing set.

($i$.) Regarding $P$ as a thin category, prove that the product $p = \prod_{i\in\lambda} a_i$, when it exists, is the supremum of $A$.
*Hint:* Recall that, if $X$ is a preorder, the **supremum** of a set $S \subseteq X$ is the element $s \in X$ such that if $a_i \leq s'$ for all $i \in \lambda$, then $s \leq s'$.

($ii$.) We know that the dual of the product is the coproduct. Can you guess what the coproduct $\coprod_{i\in\lambda} a_i$ in $P$ is in this case? Prove it.

**4.** Let $\mathcal{C}$ and $\mathcal{D}$ be categories. Consider the functor category $\mathbf{Fun}(\mathcal{C}, \mathcal{D})$. What is a product in this category? What conditions do we need to place on $\mathcal{C}$ and $\mathcal{D}$ for this product to exist?

## 3.6 Introduction to Limits and Colimits

In our previous work, we learned a lot about universal morphisms and then studied the basics of how products and coproducts behave in categories. Such studying provides a great deal of preparation for the concepts of limits and colimits, which we will introduce in this section. Before we do so, it will be convenient to utilize the notion of a *cone*.

**Definition 3.6.1.** Let $\mathcal{C}$ be a category, $A$ an object of $\mathcal{C}$. Let $F : J \longrightarrow \mathcal{C}$ be a functor, $J$ an arbitrary category. A **cone with** $A$ **over** $F$ is a family of morphisms

$$\varphi_i : A \longrightarrow F(i) \qquad i \in J$$

such that, for each morphism $f : i \longrightarrow j$ in $J$, the diagram below commutes.

$$\begin{tikzcd} & A \arrow[dl, "\varphi_i"'] \arrow[dr, "\varphi_j"] & \\ F(i) \arrow[rr, "F(f)"'] & & F(j) \end{tikzcd}$$

We denote the set of cones over $F$ with apex $A$ as $\text{Cone}(A, F)$.

Dually, a **cone with** $F$ **over** $A$ is a family of morphisms

$$\varphi_i : F(i) \longrightarrow A \qquad i \in J$$

such that, for each morphism $f : i \longrightarrow j$ in $J$, the diagram below commutes.

$$\begin{tikzcd} F(i) \arrow[rr, "F(f)"] \arrow[dr, "\varphi_i"'] & & F(j) \arrow[dl, "\varphi_j"] \\ & A & \end{tikzcd}$$

Similarly, we define the set of cones with $F$ over $A$ as $\text{Cone}(F, A)$.

We will see that the above concept is similar to the work we have done so far. To demonstrate this, we generalize our concept of a diagonal functor.

**Definition 3.6.2.** Let $\mathcal{C}$ and $J$ be categories. The **diagonal functor on** $J$ is the functor $\Delta : \mathcal{C} \longrightarrow \text{Fun}(J, \mathcal{C})$ which sends an object $C$ to the functor $\Delta(C) : J \longrightarrow \mathcal{C}$, defined as follows: Each $i \in J$ is mapped to $C$, and every morphism in $J$ is mapped to the identity of $\mathcal{C}$.

Note how if we set $J = \mathcal{D}_n$, the discrete category on $n$-object, or $J = \mathcal{D}_\lambda$, the discrete category with objects indexed by $\lambda$, we obtain our original definitions of the diagonal functor.

**Proposition 3.6.3.** Let $\mathcal{C}$ and $J$ be categories. Suppose $F : J \longrightarrow \mathcal{C}$ is a functor, and let $A$ be an object of $\mathcal{C}$.

- A cone with $A$ over $F$ corresponds to a natural transformation $\varphi : \Delta(A) \longrightarrow F$, and vice versa. In other words,
$$\text{Cone}(A, F) \cong \text{Nat}(\Delta(A), F).$$

- A cone with $F$ over $A$ corresponds to a natural transformation $\varphi : F \longrightarrow \Delta(A)$, and vice versa. In other words,
$$\text{Cone}(F, A) \cong \text{Nat}(F, \Delta(A)).$$

The proof is left to the reader. The proposition is the key to mentally switching back and forth from thinking about cones and natural transformations (between suitable functors) as equivalent constructions.

We now define limits and colimits.

**Definition 3.6.4** (Limits)**.** Let $F : J \longrightarrow \mathcal{C}$ be a functor. The **limit of** $F$ is an object $\text{Lim}\, F$ equipped with a natural transformation $u : \Delta(\text{Lim}\, F) \longrightarrow F$ such that

$$(\text{Lim}\, F, u : \Delta(\text{Lim}\, F) \longrightarrow F) \text{ is universal from } \Delta \text{ to } \text{Lim}\, F.$$

- This means that, for any other pair $(C, v : \Delta(C) \longrightarrow F)$ with $v$ a natural transformation and with $C \in \mathcal{C}$, there exists a unique morphism $h : C \longrightarrow \text{Lim}\, F$ in $\mathcal{C}$ such that the diagram below commutes.

$$\begin{tikzcd}
\Delta(\text{Lim}\, F) \arrow[r, "v"] & F \\
\Delta(C) \arrow[u, dashed, "\Delta(h)"] \arrow[ur, "v"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
\text{Lim}\, F \\
C \arrow[u, dashed, "h"']
\end{tikzcd}$$

- By Proposition 3.6.3, the morphism $u : \Delta(\text{Lim}\, F) \longrightarrow F$ forms a cone with $\text{Lim}\, F$ over $F$ via a family of morphisms $u_i : \text{Lim}\, F \longrightarrow F(i)$ for all $i \in J$.

  Similarly, any other pair $(C, v : \Delta(C) \longrightarrow F)$ is also a cone with $C$ over $F$ via a family of morphisms $v_i : C \longrightarrow F(i)$ with $i \in J$.

  Thus, the universal property, states that there exists a unique $h : C \longrightarrow \text{Lim}\, F$ such that the diagram below commutes.

$$\begin{tikzcd}
& C \arrow[d, dashed, "h"] \arrow[ddl, bend right, "v_i"'] \arrow[ddr, bend left, "v_j"] & \\
& \text{Lim}\, F \arrow[dl, "u_i"'] \arrow[dr, "u_j"] & \\
F(i) \arrow[rr, "F(f:i \longrightarrow j)"'] & & F(j)
\end{tikzcd}$$

**Remark 3.6.5.** We remind the reader that limits do not always exist for certain functors. This is because universal objects do not always exist. We will eventually discuss conditions for existence of limits.

Next, we offer the definition of a limit.

**Definition 3.6.6 (Colimits).** Let $F : J \longrightarrow \mathcal{C}$ be a functor. The **colimit of** $F$ is an object $\text{Colim}\, F$ equipped with a natural transformation $u : F \longrightarrow \Delta(\text{Colim}\, F)$ such that

$$(\text{Colim}\, F, u : F \longrightarrow \Delta(\text{Colim}\, F)) \text{ is universal from } F \text{ to } \Delta.$$

Now is a good time to use Proposition 3.6.3 and reinterpret the definition of a colimit as a family of morphisms like we did in the definition of a limit.

**Remark 3.6.7.** We comment on the notation of a limit.

- Many people denote the limit of a functor as $\text{Lim}_{\longleftarrow} F$.
- Many people denote the colimit of a functor as $\text{Lim}_{\longrightarrow} F$.

The notation makes only sense if one understand the connection between limits and colimits and universal morphisms. (Compare the direction of the arrow $h$ in the universal diagrams).

However, this then sometimes leads people to start writing $\text{Colim}_{\longrightarrow} F$ and $\text{Colim}_{\longleftarrow} F$. The issue with this notation is that it seems unnecessarily complicated (perhaps I am wrong, but I have waited for a long time to come upon an instance for when it could be useful). Despite these observations, this notation is very consistently used in texts which use categorical tools, and so this warrants a comment to the reader.

Moving forward, I will simply write $\text{Lim}\, F$ and $\text{Colim}\, F$, since I see no need to make the notation anymore complicated than it needs to be.

---

**Example 3.6.8.** Let $J = \mathcal{D}_n$, the discrete category with $n$-objects. Let $F : J \longrightarrow \mathcal{C}$ be the functor where $F(\bullet_i) = A_i$. We then have that

- The product $\prod_{i=1}^{n} A_i$ is the limit of $F$.
- The coproduct $\coprod_{i=1}^{n} A_i$ is the colimit of $F$.

When we set $J = \mathcal{D}_\lambda$, with $\lambda$ an arbitrary indexing set, we similarly get that the arbitrary product and coproduct definitions are simply instances of limits and colimits.

Thus, universal diagrams and limits have been right in our faces for the last three sections.

---

Since limits and colimits are universal objects, we have the following proposition. This is a genearlization of Proposition 3.5.8.

**Proposition 3.6.9.** Let $F : J \longrightarrow \mathcal{C}$ be a functor.

- If $\text{Lim}\, F$ exists, then for each object $C$ of $\mathcal{C}$, we have the natural bijection

$$\text{Hom}_{\mathcal{C}}(C, \text{Lim}\, F) \cong \text{Cone}(C, F)$$

- If $\operatorname{Colim} F$ exists, then for each object $C$ of $\mathcal{C}$, we have the natural bijection

$$\operatorname{Hom}_{\mathcal{C}}(\operatorname{Colim} F, C) \cong \operatorname{Cone}(F, C)$$

*Proof.* We prove the first result. Since $\operatorname{Lim} F$ exists, let $(\operatorname{Lim} F, u : \Delta(\operatorname{Lim} F) \longrightarrow F)$ be universal from $\Delta$ to $F$. Then by Exercise 3.2.1, we have the natural bijection

$$\operatorname{Hom}_{\mathcal{C}}(C, \operatorname{Lim} F) \cong \operatorname{Hom}_{\mathbf{Fun}(J,\mathcal{C})}(\Delta(C), F) = \operatorname{Nat}(\Delta(C), F).$$

By Proposition 3.6.3, we can rewrite this natural bijection as

$$\operatorname{Hom}_{\mathcal{C}}(C, \operatorname{Lim} F) \cong \operatorname{Cone}(C, F).$$

This proves the first result; the second follows similarly. ■

The above proposition is very useful as it gives us the following proposition, which is our first test of whether or not a limit or colimit exists in a category.

**Proposition 3.6.10.** Let $F : J \longrightarrow \mathcal{C}$ be a functor. Then we may define the functors

$$\operatorname{Cone}(-, F) : \mathcal{C} \longrightarrow \mathbf{Set}$$
$$\operatorname{Cone}(F, -) : \mathcal{C} \longrightarrow \mathbf{Set}$$

We have the following two results.

- $\operatorname{Cone}(-, F)$ is representable if and only if $\operatorname{Lim} F$ exists in $\mathcal{C}$ (in which case, this is the representing object)
- $\operatorname{Cone}(F, -)$ is representable if and only if $\operatorname{Colim} F$ exist in $\mathcal{C}$ (in which case, this is the representing object)

*Proof.* For pedagogical reasons, we prove the second bullet point and leave the first as an exercise.

One direction is immediate: If $\operatorname{Colim} F$ exists, then by Proposition 3.6.9, we obtain a natural bijection for each $C$ in $\mathcal{C}$ which implies that $\operatorname{Cone}(F, -)$ is representable.

Conversely, suppose $\operatorname{Cone}(F, -)$ is a representable functor with representing object $R$. We want to show $R = \operatorname{Lim} F$. Now by Proposition 3.2.6, $\operatorname{Cone}(F, -)$ is representable if and only if $(R, u : \{\bullet\} \longrightarrow \operatorname{Cone}(F, R))$ is universal from $\{\bullet\}$ to $\operatorname{Cone}(F, -)$.

Let us shut off our brains and blindly expand what this means. This means that for any other pair $(C, v : \{\bullet\} \longrightarrow \operatorname{Cone}(F, C))$, there exists a unique $h : R \longrightarrow C$ such that diagram below commutes.

$$\begin{tikzcd}
\{\bullet\} \arrow[r, "u"] \arrow[dr, "v"'] & \operatorname{Cone}(F, R) \arrow[d, dashed, "\operatorname{Cone}(F,h)"] & R \arrow[d, dashed, "h"] \\
& \operatorname{Cone}(F, C) & C
\end{tikzcd}$$

Now let us turn our brains back on and understand what this means. A function $u : \{\bullet\} \longrightarrow \text{Cone}(F, R)$ simply picks out some cone $\sigma \in \text{Cone}(F, R)$ whose family we denote as $\sigma_i : F(i) \longrightarrow R$.

Similarly, $v : \{\bullet\} \longrightarrow \text{Cone}(F, C)$ picks out a cone $v(\{\bullet\})$, which we may denote as $\tau$. What the universal property then says is the following: Given any cone $\tau$ with $F$ over some object $C$, there exists a unique $h$ such that the diagram below commutes.

$$\begin{tikzcd}
& C & \\
& R \arrow[u, dashed, "h"'] & \\
F(i) \arrow[uur, bend left, "\tau_i"] \arrow[ur, "\sigma_i"] \arrow[rr, "F(f:i\longrightarrow j)"'] & & F(j) \arrow[ul, "\sigma_j"'] \arrow[uul, bend right, "\tau_j"']
\end{tikzcd}$$

This then means that $R = \text{Colim}\, F$, which proves this direction. ∎

**Remark 3.6.11.** The above theorem is actually quite remarkable. We have linked the existence of our limit to the representability of a particular functor (one which we understand fairly well). This tells us the concept of a cone is very intimately linked to that of a limit and colimit.

## Exercises

**1.** Let $F, G : J \longrightarrow \mathcal{C}$ be two functors, and suppose $F \cong G$ (i.e., there is a natural isomorphism between them). Show that

   ($i$.) If $\text{Lim}\, F$ exists, then $\text{Lim}\, G$ exists and $\text{Lim}\, F \cong \text{Lim}\, G$.

   ($ii$.) If $\text{Colim}\, F$ exists, then $\text{Colim}\, G$ exists and $\text{Colim}\, F \cong \text{Colim}\, G$.

   Thus, limits and colimits are invariant up to isomorphism.

**2.** Prove Proposition 3.6.3.

**3.** Expand Definition 3.6.6, the definition of a colimit, in a similar fashion to how we expanded Definition 3.6.4, the definition of a limit.

**4.** Use Proposition 3.6.3 and Proposition 3.2.1 to show that if $\text{Colim}\, F$ exists for a functor $F : J \longrightarrow \mathcal{C}$, then we have a natural bijection

$$\text{Hom}_{\mathcal{C}}(\text{Colim}\, F, C) \cong \text{Cone}(F, C).$$

   This then completes the proof of Proposition 3.6.9.

**5.** Use Proposition 3.6.9 (the proof of which you just completed) to prove the first bullet point of Proposition 3.2.6: The functor $\mathrm{Cone}(-, F) : \mathcal{C} \longrightarrow \mathbf{Set}$ is representable if and only if $\mathrm{Lim}\, F$ exists. Use the following steps.

($i$.) Let $\mathcal{C}$ be a category, $F : J \longrightarrow \mathcal{C}$ a functor. Recall that we may define the functor

$$\overline{F} : J \longrightarrow \mathcal{C}^{\mathrm{op}}$$

which acts the same as $F$ on objects, but if $f : i \longrightarrow j$ is a morphism in $J$, then $\overline{F}(f) = F(f)^{\mathrm{op}}$.

Show that $\mathrm{Lim}\, F$ exists in $\mathcal{C}$ if and only if $\mathrm{Colim}\, \overline{F}$ exists in $\mathcal{C}^{\mathrm{op}}$.

$ii$. Show that

$$\mathrm{Cone}(-, F) \cong \mathrm{Cone}(\overline{F}, -).$$

Then use ($i$) and the second bullet point of Proposition 3.2.6 to complete the proof.

## 3.7 Equalizers and Coequalizers

We introduce equalizers and coequalizers as further examples of limits, and therefore examples of universal morphisms. Equalizers and coequalizers are important constructions that are useful for proofs and definitions that we will encounter later on. We first introduce examples of equalizers and coequalizers.

---

**Example 3.7.1.** Let $G$ and $H$ be groups, and consider a pair of homomorphisms $\varphi$ and $\psi$ as below.

$$\begin{tikzcd} G \arrow[r, shift left, "\varphi"] \arrow[r, shift right, "\psi"'] & H \end{tikzcd}$$

Now consider the homomorphism $\varphi - \psi : G \longrightarrow H$. Then observe that

$$\mathrm{Ker}(\varphi - \psi) = \left\{ g \in G \;\middle|\; (\varphi - \psi)(g) = 0 \right\}$$

and note that this is also the set of all $g \in G$ in which $\varphi$ and $\psi$ agree. In fact, it is the smallest such set, a notion we can make precise by the following observation: If $G'$ is another group with $\vartheta : G' \longrightarrow G$ another map such that $\varphi \circ \vartheta = \psi \circ \vartheta$, then there exists a unique $i : G' \longrightarrow \mathrm{Ker}(\varphi - \psi)$ such that the diagram below commutes.

$$\begin{tikzcd} \mathrm{Ker}(\varphi - \psi) \arrow[r, "i"] & G \arrow[r, shift left, "\varphi"] \arrow[r, shift right, "\psi"'] & H \\ G' \arrow[u, dashed, "i'"] \arrow[ur, "\vartheta"'] & & \end{tikzcd}$$

Note above that $i : \mathrm{Ker}(\varphi - \psi) \longrightarrow G$ is the inclusion morphism. Also note that this construction is possible for any two parallel group homomorphisms.

---

**Example 3.7.2.** In **Set**, equalizers always exist. Simply let $D = \{x \in A \mid f(x) = g(x)\}$, and let $e : D \longrightarrow A$ by the inclusion morphism into $A$. Clearly we'll have that $f \circ e = g \circ e$.

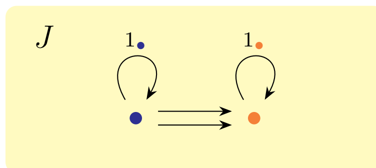
Now for any $h : C \longrightarrow A$ such that $f \circ h = g \circ h$, we see that the image of $h$ must be a subset of $D$. Hence there exists a unique inclusion morphism $i : C \longrightarrow D$, which shows that $e$ in fact is the equalizer in **Set** for any $f, g : A \longrightarrow B$.

---

**Definition 3.7.3** (Nice Equalizer Definition)**.** Let $\mathcal{C}$ be a category and consider a pair of parallel morphisms $f, g : A \longrightarrow B$. The equalizer of $f$ and $g$ is a pair $(E, e : E \longrightarrow A)$ such that $f \circ e = g \circ e$ with the following property. For any other morphism $h : C \longrightarrow A$ such that $f \circ h = g \circ h$, there exists a unique morphism $f' : C \longrightarrow E$ such that the following commutes.

$$\begin{tikzcd} E \arrow[r, "e"] & A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \\ C \arrow[u, dashed, "f'"] \arrow[ur, "h"'] & & \end{tikzcd}$$

**Definition 3.7.4** (Equalizer as a Limit)**.** Let $\mathcal{C}$ be a category and consider a pair of parallel morphisms $f, g : A \longrightarrow B$. Let $J$ be the category with two elements and two nontrivial morphisms as below.




and let $F : J \longrightarrow \mathcal{C}$ be the functor such that $F(\bullet \rightrightarrows \bullet) = A \overset{f}{\underset{g}{\rightrightarrows}} B$ We define the **equalizer** of $f$ and $g$ to be limit $(\text{Lim}\, F, e : \Delta(\text{Lim}\, F) \longrightarrow F)$ of $F$.

**Proposition 3.7.5.** Let $\mathcal{C}$ be a category, and suppose $e : D \longrightarrow A$ is an equalizer for a pair of morphisms $f, g : A \longrightarrow B$. Then $e$ is monic.

*Proof.* Consider any pair $f_1, f_2 : C \longrightarrow D$ such that $e \circ f_1 = e \circ f_2$. Then we have that

$$\begin{tikzcd} C \arrow[r, shift left, "f_1"] \arrow[r, shift right, "f_2"'] & D \arrow[r, "e"] & A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \end{tikzcd}$$

Since $e \circ f_1 = e \circ f_2$, we see that

$$\begin{aligned} f \circ e = g \circ e &\implies f \circ (e \circ f_1) = g \circ (e \circ f_1) \\ &\implies f \circ (e \circ f_1) = g \circ (e \circ f_2). \end{aligned}$$

Hence we see $e \circ f_1 = e \circ f_2 : C \longrightarrow D$ is another morphism which is equalized by $f$ and $g$.

$$\begin{tikzcd} D \arrow[r, "e"] & A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \\ C \arrow[u, dashed, "f'"] \arrow[ur, "e \circ f_1 = e \circ f_2"'] & & \end{tikzcd}$$

By the universality of the equalizer $e : D \longrightarrow A$, we know that there must exist a unique morphism $f' : C \longrightarrow D$ such that

$$e \circ f' = e \circ f_1 = e \circ f_2.$$

Since $f'$ is unique, we are forced to conclude that $f_1 = f_2$. Hence $e \circ f_1 = e \circ f_2 \implies f_1 = f_2$, so that $e : D \longrightarrow A$ is monic. ∎

**Definition 3.7.6.** Let $\mathcal{C}$ be a category with a zero object $Z$ of $\mathcal{C}$. That is, an object which is both initial and terminal, such that for any objects $A, B$ of $\mathcal{C}$ there exists a unique pair of morphisms $f, g$ such that

$$A \xrightarrow{f} Z \xrightarrow{g} B.$$

Denote $f \circ g = 0$ as the zero arrow (any morphism which passes through $z$ is a zero arrow).

Now we define the **cokernel** a morphism $f : A \longrightarrow B$ to be an arrow $u : B \longrightarrow C$ where

1. $u \circ f = 0 : A \longrightarrow C$
2. If $h : B \longrightarrow D$ has the property that $h \circ f = 0$, then $h = h' \circ u$ for a unique arrow $h' : B \longrightarrow D$.

Visually, this becomes

$$\begin{tikzcd} A \arrow[r, "f"] & B \arrow[r, "u"] \arrow[dr, "h"'] & E \arrow[d, dashed, "h'"] \\ & & C. \end{tikzcd}$$

The cokernel is a special object in **Ab**, as it plays a role in the concept of exact sequences and hence homology as well. The cokernel of a homomorphism $f : G \longrightarrow H$ is the projection $H \longrightarrow H/\operatorname{Im}(G)$, a quotient group of $B$. This is often written as

$$\operatorname{coker}(f) = H/\operatorname{Im}(G).$$

## Coequalizers.

**Definition 3.7.7.** Let $\mathcal{C}$ be a category and consider two morphisms $f, g : A \longrightarrow B$ in $\mathcal{C}$. The **coequalizer** of $(f, g)$ is a morphism $u : B \longrightarrow D$ such that

1. $u \circ f = u \circ h$
2. If $h : B \longrightarrow C$ has the property that $h \circ f = h \circ g$, then there exists a unique morphism $h' : D \longrightarrow C$ such that $h = h' \circ u$.

This may not always exist. We can represent this with the following commutative diagram. Note that we can interpret a coequalizers as a morphism which uniquely "flattens" morphisms, and for any other morphism which also "flattens" is related to the original coequalizer.

$$\begin{tikzcd} A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] & B \arrow[r, "u"] \arrow[dr, "h"'] & D \arrow[d, dashed, "h'"] \\ & & C \end{tikzcd}$$

With coequalizers, we get the following nice result.

**Lemma 3.7.8.** All coequalizers are epimorphisms.

Coequalizers can also be realized as universal arrows. First consider the category $\mathbf{2}$, containing two objects and two nontrivial morphisms. Since there are only two objects, the two nontrivial morphisms have the same domain and codomain. Now consider the functor category $\mathcal{C}^{\mathbf{2}}$ where

1. Objects are functors $F : \mathbf{2} \longrightarrow \mathcal{C}$, whose image is therefore a pair of morphism $f, g : A \longrightarrow B$ in $\mathcal{C}$
2. Morphisms are natural transformations, which are therefore a pair of arrows $h : A \longrightarrow A'$ and $k : B \longrightarrow B'$ so that

$$\begin{tikzcd}
A \arrow[r, shift left, "f"] \arrow[r, shift right, "g"'] \arrow[d, "h"'] & B \arrow[d, "k"] \\
A' \arrow[r, shift left, "f'"] \arrow[r, shift right, "g'"'] & B'
\end{tikzcd}$$

is a commutative diagram. Finally consider the diagonal functor $\Delta : \mathcal{C} \longrightarrow \mathcal{C}^{\mathbf{2}}$ where

$$\begin{aligned}
C &\longmapsto (1_C, 1_C) \\
r : C \longrightarrow C' &\longmapsto (r, r).
\end{aligned}$$

Now consider a pair $f, g : A \longrightarrow B$ in $\mathcal{C}^{\mathbf{2}}$. If we have a morphism $h : B \longrightarrow C$ such that $h \circ f = h \circ g$, then this is the same thing as a morphism $(hf, hg) : (f, g) \longrightarrow (1_C, 1_C)$ in $\mathcal{C}^{\mathbf{2}}$. Therefore a coequalizer $u : B \longrightarrow C$ is a universal arrow from $(f, g)$ to $\Delta$.

---

**Example 3.7.9.** In the category $\mathbf{Ab}$, the coequalizer of two group homomorphisms $\varphi, \psi : G \longrightarrow H$ is the homomorphism

$$\pi : H \longrightarrow H/\operatorname{Im}(\varphi - \psi).$$

where $g' \in H$ maps to the coset $g' + \operatorname{Im}(\varphi - \psi)$. We show this as follows.

$\boldsymbol{\pi \circ \varphi = \pi \circ \psi}$**.** First let $g \in G$, and consider the elements

$$\begin{aligned}
\pi \circ \varphi(g) &= \varphi(g) + \operatorname{Im}(\varphi - \psi) \\
\pi \circ \psi(g) &= \psi(g) + \operatorname{Im}(\varphi - \psi).
\end{aligned}$$

If we subtract these two quantities, we get that

$$\begin{aligned}\pi \circ \varphi(g) - \pi \circ \psi(g) &= \Big[\varphi(g) + \operatorname{Im}(\varphi - \psi)\Big] - \Big[\psi(g) + \operatorname{Im}(\varphi - \psi)\Big]\\ &= (\varphi(g) - \psi(g)) + \operatorname{Im}(\varphi - \psi)\\ &= 0 + \operatorname{Im}(\varphi - \psi).\end{aligned}$$

Since their difference is zero, we see that they're equal. Hence $\pi \circ \varphi = \pi \circ \psi$.

**Universality.** Let $f : H \longrightarrow H'$ be another group homomorphism such that $f \circ \varphi = f \circ \psi$. Then construct the morphism $f' : H/\operatorname{Im}(\varphi - \psi) \longrightarrow H'$ where

$$h + \operatorname{Im}(\varphi - \psi) \longmapsto f(h).$$

Clearly this is well defined, since if $h + \operatorname{Im}(\varphi - \psi) = h' + \operatorname{Im}(\varphi - \psi)$, then this means that $h = h' + (\varphi - \psi)(g)$, so that

$$\begin{aligned}f'(h + \operatorname{Im}(\varphi - \psi)) &= f(h)\\ &= f(h' + \varphi(g) - \psi(g))\\ &= f(h') + f \circ \varphi(g) - f \circ \psi(g)\\ &= f(h')\end{aligned}$$

where in the last step we used the fact that $f \circ \varphi = f \circ \psi$. Thus we see that $f'$ is a well-defined group homomorphism. Furthermore, note that $f = f' \circ \pi$. To finally show that $f'$ is unique, we suppose there exists another group homomorphism $k : H/\operatorname{Im}(\varphi - \psi) \longrightarrow H'$ such that $f = k \circ \pi$. Then we see that $f' \circ \pi = k \circ \pi$, which implies that $f' = k$.

What we've shown is that for any $f : H \longrightarrow H'$ such that $f \circ \varphi = f \circ \psi$, there exists a unique morphism $f' : H/\operatorname{Im}(\varphi - \psi) \longrightarrow H'$ such that $f = f' \circ \pi$. Thus we see that $\pi$ has the universal property of being a coequalizer.

# 3.8 Pullbacks and Pushouts

## Pullbacks.

**Definition 3.8.1.** Let $f : A \longrightarrow C$ and $g : B \longrightarrow C$ be two morphisms. Then we say a pullback of $f, g$ is a commutative square on the left

$$
\begin{tikzcd}
D \arrow[r, "h"] \arrow[d, "k"'] & A \arrow[d, "f"] \\
B \arrow[r, "g"'] & C
\end{tikzcd}
\qquad
\begin{tikzcd}
E \arrow[r, "h'"] \arrow[d, "k;"'] & A \arrow[d, "f"] \\
B \arrow[r, "g"'] & C
\end{tikzcd}
\qquad
\begin{tikzcd}
E \arrow[dr, dashed, "f'"] \arrow[drr, bend left, "h"] \arrow[ddr, bend right, "k"'] & & \\
& D \arrow[r, "h'"] \arrow[d, "k'"'] & A \arrow[d, "f"] \\
& B \arrow[r, "g"'] & C
\end{tikzcd}
$$

such that for any commutative square in the the middle, the diagram on the right commutes, and $f'$ is unique.

Another way we can describe this is using the language of limits, and hence show that pullbacks are simply limit objects. Let $J$ be the category of three objects with the following shape:

$$1 \longrightarrow 2 \longleftarrow 3$$

The numbers 1, 2, and 3 here mean nothing; they are simply place holders for *some* distinct objects. So any functor $F : J \longrightarrow \mathcal{C}$ simply corresponds to a triple of object and a pair of morphisms in $\mathcal{C}$:

$$A \xrightarrow{f} C \xleftarrow{g} B.$$

if we have $F(1) = A$, $F(2) = C$ and $F(3) = B$. Now we can equivalently describe a pullback as follows:

**Definition 3.8.2.** If $J$ is the category with the shape $1 \longrightarrow 2 \longleftarrow 3$ , and $F : J \longrightarrow \mathcal{C}$ is a functor, then a **pullback** is a universal arrow $(D, u : \Delta(D) \longrightarrow F)$ from $\Delta$ to $F$.

First, observe that this shows that a pullback is a limit. But how are our two definitions equivalent?

Consider the morphism $u : \Delta(D) \longrightarrow F$. This is simply a natural transformation between the two functors $\Delta(D) : J \longrightarrow \mathcal{C}$ and $F : J \longrightarrow \mathcal{C}$. Now $\Delta(D)(i) = D$ for all objects $i = 1, 2, 3 \in J$. On the other hand, $F(1) = A$, $F(2) = C$ and $F(3) = B$. Thus we see that $\Delta(R) \longrightarrow F$ induces a family of morphisms:

$$
\begin{aligned}
u_1 : \Delta(D)(1) \longrightarrow F(1) &\implies u_1 : D \longrightarrow A \\
u_2 : \Delta(D)(2) \longrightarrow F(2) &\implies u_2 : D \longrightarrow C \\
u_3 : \Delta(D)(3) \longrightarrow F(3) &\implies u_3 : D \longrightarrow B
\end{aligned}
$$

which arrange themselves in $\mathcal{C}$ into the following diagram:

$$
\begin{tikzcd}
 & D \arrow[dl, "u_1"'] \arrow[d, "u_2"] \arrow[dr, "u_3"] & \\
A \arrow[r, "f"'] & C & B \arrow[l, "g"]
\end{tikzcd}
$$

and if we "tip" this diagram over, and force the arrows $f$ and $g$ meeting at $C$ into a 90 degree angle, we get the following cone:

$$
\begin{tikzcd}
D \arrow[r, "u_1"] \arrow[d, "u_3"'] \arrow[dr, "u_2"] & A \arrow[d, "f"] \\
B \arrow[r, "g"'] & C
\end{tikzcd}
=
\begin{tikzcd}
D \arrow[r, "u_1"] \arrow[d, "u_3"'] & A \arrow[d, "f"] \\
B \arrow[r, "g"'] & C
\end{tikzcd}
$$

Note that we removed the morphism $u_2$ because it's redundant, unnecessary information; after all $u_2 = f \circ u_1 = g \circ u_3$; which is information already captured in both the original diagram and the commutative square.

Thus, we see that whenever we have an object $E$ and morphism $v : \Delta(E) \longrightarrow F$, we have a commutative square! In other words, whenever we have a cone over $F$, we have a commutative square! And in even *other* words, whenever we have a family of morphisms $v_i : E \longrightarrow F(i)$ for $i = 1, 2, 3$, we have a commutative square!

$$
\begin{tikzcd}
E \arrow[r, "v_1"] \arrow[d, "v_3"'] \arrow[dr, "v_2"] & A \arrow[d, "f"] \\
B \arrow[r, "g"'] & C
\end{tikzcd}
=
\begin{tikzcd}
E \arrow[r, "v_1"] \arrow[d, "v_3"'] & A \arrow[d, "f"] \\
B \arrow[r, "g"'] & C
\end{tikzcd}
$$

So, how do we connect the universality of $(D, u : \Delta(D) \longrightarrow F)$ with the universality of the pullback? Well, since this object is universal, we know that for any other pair $(E, v : \Delta(E) \longrightarrow F)$, there exists a morphism $f' : E \longrightarrow D$ such that the following diagram commutes.

$$
\begin{tikzcd}
F & \Delta(D) \arrow[l, "u"'] \\
 & \Delta(E) \arrow[ul, "v"] \arrow[u, dashed, "\Delta(f')"']
\end{tikzcd}
\qquad
\begin{tikzcd}
D \\
E \arrow[u, dashed, "f'"]
\end{tikzcd}
$$

The commutativity of the top left diagram gives us the relation that $u \circ \Delta(f') = v$, which implies that $u_1 \circ f' = v_1$ and $u_3 \circ f' = v_3$. We then have that

$$
\begin{tikzcd}
E \arrow[r, "v_1"] \arrow[d, "v_3"'] & A \arrow[d, "f"] \\
B \arrow[r, "g"'] & C
\end{tikzcd}
=
\begin{tikzcd}
E \arrow[r, "u_1 \circ f'"] \arrow[d, "u_3 \circ f'"'] & A \arrow[d, "f"] \\
B \arrow[r, "g"'] & C
\end{tikzcd}
=
\begin{tikzcd}
E \arrow[dr, dashed, "f'"] \arrow[drr, bend left, "v_1"] \arrow[ddr, bend right, "v_2"'] & & \\
& D \arrow[r, "u_1"] \arrow[d, "u_3"'] & A \arrow[d, "f"] \\
& B \arrow[r, "g"'] & C
\end{tikzcd}.
$$

which is just the pullback. Thus the pullback is in fact a limit object, and we understand just exactly how it is a limit object of the functor $F : J \longrightarrow \mathcal{C}$.

**Definition 3.8.3.** Let $\mathcal{C}$ be a category, and consider a pair of morphism $f : A \longrightarrow B$, $g : A \longrightarrow C$ in $\mathcal{C}$. A **pushout** of $(f, g)$ is the commutative diagram on the left

$$
\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "g"'] & B \arrow[d, "u"] \\
C \arrow[r, "v"'] & R
\end{tikzcd}
\qquad
\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "g"'] & B \arrow[d, "h"] \\
C \arrow[r, "k"'] & S
\end{tikzcd}
$$

such that for every commutative square as on the right, there exists a unique morphism $t : R \longrightarrow S$ such that $t \circ u = h$ and $t \circ v = k$. We can actually summarize this information more compactly

$$
\begin{tikzcd}
A \arrow[r, "f"] \arrow[d, "g"'] & B \arrow[d, "u"] \arrow[ddr, bend left, "h"] & \\
C \arrow[r, "v"'] \arrow[drr, bend right, "k"'] & R \arrow[dr, dashed, "t"] & \\
& & S
\end{tikzcd}
$$

where the diagram is commutative. One way to imagine a pushout is a commutative diagram which swallows every other commutative diagram which contains the morphisms $f, g$.

As you might suspect, the pushout can in fact be related as the universal arrow of a functor. Consider the category $\mathbf{3}$, which contains 3 objects and two nontrivial morphisms.

$$
Y \xleftarrow{f} X \xrightarrow{g} Z
$$

Now construct the functor category $\mathcal{C}^{\mathbf{3}}$, where

1. Objects are functors $F : \mathbf{3} \longrightarrow \mathcal{C}$, which is equivalent to pairs of morphisms $(f, g)$ where $f : A \longrightarrow B$ and $g : A \longrightarrow C$ in $\mathcal{C}$

2. Morphisms are natural transformations, which in this case simply reduce to a triple of morphisms $(h, l, k)$ where

$$\begin{tikzcd}
B \arrow[d, "h"'] & A \arrow[l, "f"'] \arrow[r, "g"] \arrow[d, "l"] & C \arrow[d, "k"] \\
B' & A' \arrow[l, "f'"] \arrow[r, "g'"'] & C'
\end{tikzcd}$$

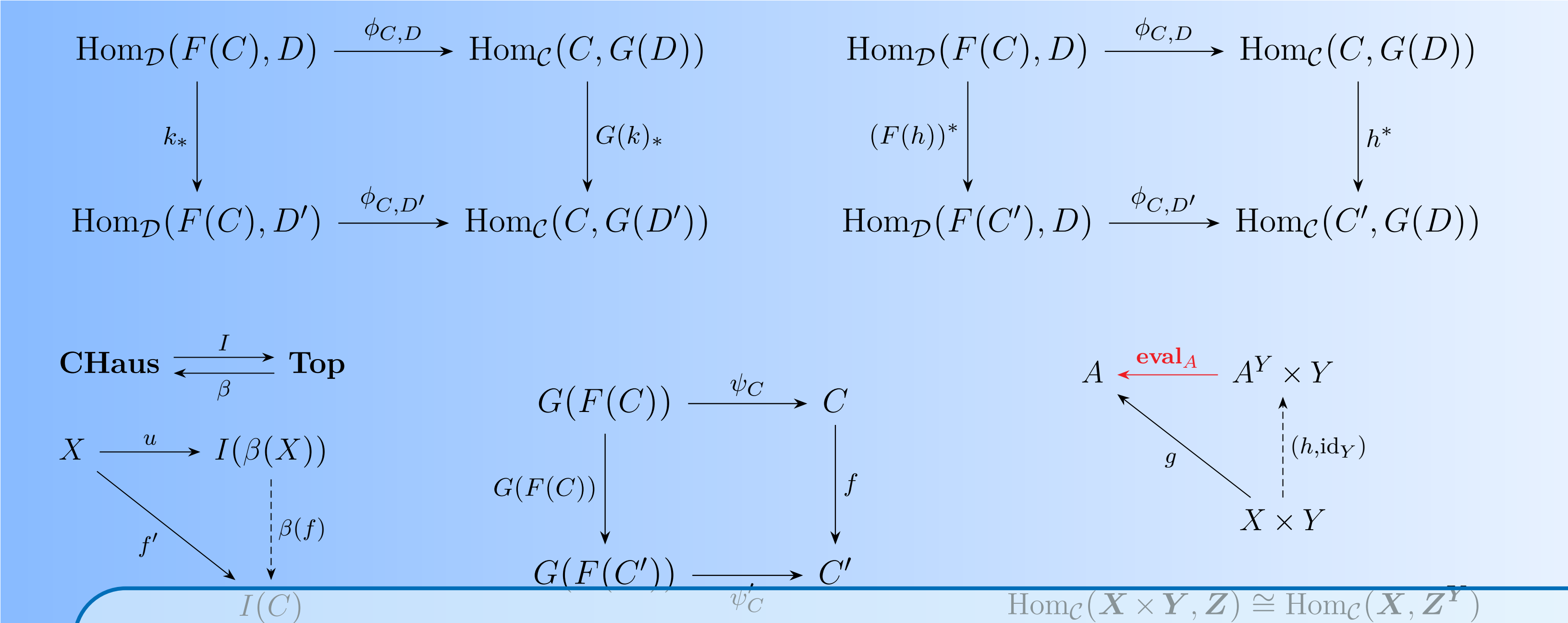
Now construct the functor $\Delta : \mathcal{C} \longrightarrow \mathcal{C}^{\mathbf{3}}$ where $C \longmapsto (1_C, 1_C)$ where $1_C : C \longrightarrow C$ is the identity morphism. Suppose there exists a natural transformation $\eta_S : (f, g) \longrightarrow \Delta(S)$, which we can represent as follows:

$$\begin{tikzcd}
B \arrow[d, "h"'] & A \arrow[l, "f"'] \arrow[r, "g"] \arrow[d, "l"] & C \arrow[d, "k"] \\
S & S \arrow[l, "1_S"] \arrow[r, "1_S"'] & S
\end{tikzcd}$$

If we have a pushout associated with the object $R$ in $\mathcal{C}$, the existence of these commutative squares implies the existence of a morphism $t : R \longrightarrow S$, so that we have

$$\begin{tikzcd}
(f, g) \arrow[r, "\eta_R"] \arrow[dr, "\eta_S"'] & \Delta(R) \arrow[d, dashed, "\Delta(t)"] & R \arrow[d, dashed, "t"] \\
& \Delta(S) & S
\end{tikzcd}$$

Hence we see that a pushout is a universal arrow from $(f, g)$ to $\Delta$.



# 4. Adjunctions.

## 4.1 Introduction to Adjunctions.

As promised, we now build upon the work we did with universal morphisms to define the concept of an adjunction. Adjunctions are special cases of universal morphisms that occur between two functors $F$ and $G$ which assemble between two categories $\mathcal{C}$ and $\mathcal{D}$ as below.

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

Studying adjunctions allows us to give an answer to many questions that appear in categories. For example, adjunctions can explain why, for instance, given two sets $X$, $Y$, we have the isomorphism

$$F(X \times Y) \cong F(X) * F(Y)$$

where $F : \textbf{Set} \longrightarrow \textbf{Grp}$ is the free group functor and $*$ denotes the free product. They can also explain why this property, and other similar properties, hold for similar free functors.

We begin with an example of an adjunction.

---

**Example 4.1.1.** Recall that for a fixed unital ring $R$ in **Ring**, we may form the functor

$$R[-] : \textbf{Grp} \longrightarrow R - \textbf{Alg}$$

which sends a group $G$ to its group ring $R[G]$. Recall that

$$R[G] = \left\{ \sum_{g \in G} a_g g \;\middle|\; g \in G,\ a_g \in R, \text{ and } a_g = 0 \text{ for all but finitely many } a_g \right\}.$$

Recall also that we can form the functor

$$(-)^\times : R\text{-}\mathbf{Alg} \longrightarrow \mathbf{Grp}$$

which sends an $R$-algebra $A$ to its group of units $A^\times$. These two functors are related in the following way. Consider a group $G$ and its group ring $R[G]$. In general, the units of $R[G]$ are nontrivial. One thing we do know is that elements of the form $1_R g$, with $g \in G$, are units of $R[G]$. (The multiplicative inverse of such an element is $1_R g^{-1}$.) This allows us to construct a group homomorphism

$$i : G \longrightarrow (R[G])^\times \qquad g \mapsto 1_R g.$$

What is interesting about this is the following fact: $(G, i : G \longrightarrow (R[G])^\times)$ is universal from $G$ to $(-)^\times$. That is, if $K$ is a ring, and we have a mapping $\varphi : G \longrightarrow K^\times$, then there exists a unique ring homomorphism $h : R[G] \longrightarrow K$ such that the diagram below commutes.

$$\begin{tikzcd} G \arrow[r, "i"] \arrow[dr, "\varphi"'] & (R[G])^\times \arrow[d, dashed, "(h)^\times"] & & R[G] \arrow[d, dashed, "h"] \\ & (K)^\times & & K \end{tikzcd}$$

The reason why this works is as follows: $\varphi$ tells us to where to send elements of $G$. Since a map on $R[G]$ can be defined by (1) defining where elements of $G$ go and (2) extending linearly, $\varphi$ induces the existence of $h$.

By Proposition 3.2.1, we then have the following result: If $K$ is an $R$-algebra, then for each group $G$ there is a natural bijection

$$\text{Hom}_{\mathbf{Ring}}(R[G], K) \cong \text{Hom}_{\mathbf{Grp}}(G, (K)^\times)$$

Specifically, the bijection is natural in $G$.

But wait—There's more! For every ring $K$, there is a natural ring homomorphism

$$\varepsilon : R[(K)^\times] \longrightarrow K \qquad \sum_{k \in K^\times} a_k k \mapsto z(a_k)k$$

where $z(a_k) = 1_k$, the identity of $K$, if $a_k \neq 0$, and $z(a_k) = 0$ if $a_k = 0$. The reason why we care about this is because $(K, R[(K)^\times] \longrightarrow K)$ is universal from $R[-]$ to $(K)^\times$. That is, if $G$ is a group and we have a mapping $\psi : R[G] \longrightarrow K$, then there exists a unique $j : G \longrightarrow (K)^\times$ such that the following diagram commutes.

$$\begin{tikzcd} R[(K)^\times] \arrow[r, "\varepsilon"] & K & & (K)^\times \\ R[G] \arrow[u, dashed, "R[j]"] \arrow[ur, "\psi"'] & & & G \arrow[u, dashed, "j"] \end{tikzcd}$$

We obtain $j$ as follows: Note that $\varphi(1_R g) \in K^\times$, since ring homomorphisms send units to units. Hence, the composite

$$G \xrightarrow{i} (R[G])^\times \xrightarrow{\varphi^\times} K^\times$$

where $i$ is defined earlier, yields $j$. Moreover, the diagram commutes in this way. By Exercise 3.2, if $K$ is a ring, then for every group $G$ we have the following natural bijection

$$\text{Hom}_{\mathbf{Grp}}(G, (K)^\times) \cong \text{Hom}_{\mathbf{Ring}}(R[G], K).$$

Specifically, the bijection is natural in $K$. However, we just saw this isomorphism before! This demonstrates our first example of an adjunction.

---

**Definition 4.1.2.** Let $\mathcal{C}, \mathcal{D}$ be categories. Consider a pair of functors

$$\begin{tikzcd} \mathcal{C} \arrow[r, shift left, "F"] & \mathcal{D} \arrow[l, shift left, "G"] \end{tikzcd}$$

We say that $F, G$ form an **adjunction** and that $F$ is **left adjoint to** $G$ (and so $G$ is **right adjoint to** $F$) if, for all $C \in \mathcal{C}$, $D \in \mathcal{D}$, there is a natural bijection

$$\text{Hom}_{\mathcal{D}}\Big(F(C), D\Big) \cong \text{Hom}_{\mathcal{C}}\Big(C, G(D)\Big)$$

This definition is somewhat strange, so we comment a few remarks.

**Remark 4.1.3.**

- To define an adjunction between two functors, it suffices to specify which functor is the left adjoint, or which functor is the right adjoint (since one specification determines the other). Thus, the sentence "$F$ and $G$ form an adjunction" alone does not make sense; namely, it is missing information of which functor is the left or the right adjoint.
- In an adjunction, we are always going to have some kind of bijection as above. But there are two different ways we could decide to write it:

  $$\text{Hom}_{\mathcal{D}}(F(C), D) \cong \text{Hom}_{\mathcal{C}}(C, G(D)) \quad \text{or} \quad \text{Hom}_{\mathcal{C}}(C, G(D)) \cong \text{Hom}_{\mathcal{D}}(F(C), D)$$

  This can potentially confuse us on which functor is the left adjoint, and which one is the right. However, one thing that does not change in the above expressions is the position of $F(C)$ and $G(D)$ in their hom-sets. In their hom-sets, the symbol $F(C)$ is always in the left position, while $G(D)$ is in the right. Hence we can determine if $F$ or $G$ is left or right based on glancing at the bijection. Conversely, knowing the left and rightedness of our functors tells us how to write down the bijection.

We now observe that this definition is equivalent to the existence of universal morphisms; this is something we already saw in our introductory example.

**Proposition 4.1.4.** Let $\mathcal{C}, \mathcal{D}$ be categories and consider a pair of functors $\mathcal{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathcal{D}$. The following are equivalent.

(*i.*) The functors $F$, $G$ form an adjunction where $F$ is left adjoint to $G$ (and so $G$ is right adjoint to $G$).

(*ii.*) There exist natural transformations

$$\eta : I_{\mathcal{C}} \longrightarrow G \circ F \qquad \varepsilon : F \circ G \longrightarrow I_{\mathcal{D}}$$

such that

- For each $C \in \mathcal{C}$, the morphism $\eta_C : C \longrightarrow G(F(C))$ is universal from $C$ to $G$
- For each $D \in \mathcal{D}$, the morphism $\varepsilon_D : F(G(D)) \longrightarrow D$ is universal from $F$ to $D$

*Proof.* Since $F$ is left adjoint to $G$, we have the natural bijection

$$\text{Hom}_{\mathcal{D}}(F(C), D) \cong \text{Hom}_{\mathcal{C}}(C, G(D)).$$

This is natural in $C$ and $D$.

By Proposition 3.2.1, the above bijection is natural in $D$ if and only if there exists a morphism $\eta_C : C \longrightarrow G(F(C))$ which is universal from $C$ to $G$. However, the bijection holds for all $C$. Therefore, we obtain a family of universal morphisms

$$\eta_C : C \longrightarrow G(F(C)).$$

Since this bijection is also natural in $C$, we ultimately obtain a natural transformation $\eta : I_{\mathcal{C}} \longrightarrow G \circ F$.

Using the same bijection from our adjunction, we can use Exercise 3.2 to conclude the existence of a family of morphisms $\varepsilon_D : F(G(D)) \longrightarrow D$ which is universal from $F$ to $D$. We then use the fact that the bijection is natural to form the natural transformation $\varepsilon : F \circ G \longrightarrow I_{\mathcal{D}}$, as desired.

As we used if and only if propositions, our work proves both directions, which completes the proof. ∎

**Definition 4.1.5.** Let $\mathcal{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathcal{D}$ be an adjunction. We establish the following terminology.

- The natural transformation $\eta : I_{\mathcal{C}} \longrightarrow G \circ F$ is the **unit of the adjunction**.
- The natural transformation $\varepsilon : F \circ G \longrightarrow I_{\mathcal{D}}$ is the **counit of the adjunction**.

---

**Example 4.1.6.** We already saw this proposition in action in the introductory example. In that example, we found a pair functors

$$\mathbf{Grp} \underset{(-)^{\times}}{\overset{R[-]}{\rightleftarrows}} \mathbf{Ring}$$

that formed an adjunction with universal morphisms

$$i_G : G \longrightarrow (R[G])^{\times} \qquad \varepsilon_K : R[(K)^{\times}] \longrightarrow K$$

for all groups $G$ and rings $K$. Hence $i_G$ is the unit of the adjunction, while $\varepsilon_K$ is the counit. These units and counits are what allowed us to establish the bijection

$$\mathrm{Hom}_{\mathbf{Ring}}(R[G], K) \cong \mathrm{Hom}_{\mathbf{Grp}}(G, (K)^{\times})$$

natural in $G$ and $K$. Hence, the group ring functor $R[-]$ is left adjoint to the group of units functor $(-)^{\times}$.

---

Using our previous work, we very quickly and (hopefully) painlessly established a connection between the natural bijection that appears in the definition of an adjunction and the unit and counit morphisms. However, we did not really describe what the bijection actually does on elements. The next proposition characterizes the bijection.

**Proposition 4.1.7.** Let $\mathcal{C}$, $\mathcal{D}$ be categories, and suppose $\mathcal{C} \underset{G}{\overset{F}{\rightleftarrows}} \mathcal{D}$ form an adjunction with $F$ left adjoint to $G$. Let $\eta$, $\varepsilon$ be the unit and counit.

For each $C, D$, the natural bijection

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

is given by the function where for each $f : F(C) \longrightarrow D$ and $g : C \longrightarrow G(D)$,

$$\varphi(f) = G(f) \circ \eta_C \qquad \varphi^{-1}(g) = \varepsilon_D \circ F(g).$$

The proof is left to the reader.

---

**Example 4.1.8.** We have already encountered the pair of functors

$$\mathbf{Set} \underset{U}{\overset{F}{\rightleftarrows}} \mathbf{Mon}$$

where $F$ is the free monoid functor and $U$ is the forgetful monoid functor. We previously saw that given a set $X$, there exists an inclusion morphism

$$i_X : X \longrightarrow U(F(X))$$

and this morphism is universal from $X$ to $U$. In addition, we know that the monoid homomorphism

$$\varepsilon_M : F(U(M)) \longrightarrow M$$

and this morphism is from $F$ to $M$. Therefore, we see that $F$ and $U$ are adjoint functors; specifically, $F$ is left adjoint to $G$ and $G$ is right adjoint to $F$, and we have the natural bijection

$$\text{Hom}_{\textbf{Mon}}(F(X), M) \cong \text{Hom}_{\textbf{Set}}(X, U(M)).$$

Moreover, we know exactly how this bijection works.

- For $f : F(X) \longrightarrow M$, we send $\varphi(f)$ to $U(f) \circ i_X$.
- For $g : X \longrightarrow U(M)$, we send $\varphi^{-1}(g)$ to $\varepsilon_M \circ F(g)$.

This data assembles into the commutative diagrams as below.

$$\begin{tikzcd}
X \arrow[r, "i_X"] \arrow[dr, "\varphi(f)"'] & U(F(X)) \arrow[d, "U(f)"] \\
& U(M)
\end{tikzcd}
\qquad
\begin{tikzcd}
F(U(M)) \arrow[r, "\varepsilon_M"] & M \\
F(X) \arrow[u, "F(g)"] \arrow[ur, "\varphi^{-1}(g)"']
\end{tikzcd}$$

---

Now we offer some sufficient conditions for establishing an adjunction.

**Proposition 4.1.9.** Let $G : \mathcal{D} \longrightarrow C$ be a functor. Suppose that for each $C \in \mathcal{C}$, there exists an object $F_0(C) \in \mathcal{D}$ and a universal morphism $\eta_C : C \longrightarrow G(F_0(C))$ from $C$ to $G$. Then there exists a functor $F : \mathcal{C} \longrightarrow \mathcal{D}$ which is left-adjoint to $G$.

*Proof.* To have universality from $C$ to $G$, the diagram

$$\begin{tikzcd}
C \arrow[r, "\eta_C"] \arrow[dr, "g"'] & G(F(C)) \arrow[d, dashed, "G(f)"] \\
& G(D)
\end{tikzcd}
\qquad
\begin{tikzcd}
F(C) \arrow[d, dashed, "f"'] \\
D
\end{tikzcd}$$

must commute. Hence we have a bijection

$$\text{Hom}_{\mathcal{D}}(F(C), D) \cong \text{Hom}_{\mathcal{C}}(C, G(D)).$$

Now suppose $h : C \longrightarrow C'$. Then the dashed arrow

$$\begin{tikzcd}
C \arrow[r, "\eta_C"] \arrow[d, "f"'] & G(F_0(C)) \arrow[d, dashed, "G(h)"] \\
C' \arrow[r, "\eta_{C'}"'] & G(F_0(C'))
\end{tikzcd}$$

must exist by universality; we simply utilize the previous diagram. In other words, if $h : C \longrightarrow C'$, then there exists a morphism $f : F_0(C) \longrightarrow F_0(C')$. With that said, we can then define a functor where $F : \mathcal{C} \longrightarrow \mathcal{D}$ with $F(C) = F_0(C)$ and $F(h) = F_0(C) \longrightarrow F_0(C')$. By construction, this functor is left adjoint to $G$. $\blacksquare$

A similar proposition holds for the establishing a right adjoint.

**Proposition 4.1.10.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor. Suppose for each object $D \in \mathcal{D}$ there exists an object $G_0(D) \in \mathcal{C}$ and a universal morphism $\varepsilon_d : F(G_0(D)) \longrightarrow D$ from $F$ to $D$.

Then there exists a functor $G : \mathcal{D} \longrightarrow \mathcal{C}$ which is right-adjoint to $F$.

We now introduce a proposition which offers sufficient conditions for an adjunction, although it is not parallel to either of our previous propositions.

**Proposition 4.1.11.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ and $G : \mathcal{D} \longrightarrow \mathcal{C}$ be functors, and suppose we have the pair of natural transformations:

$$\eta_C : I_C \longrightarrow G \circ F \quad \varepsilon_D : I_D \longrightarrow F \circ G$$

such that the following composites are the identity:

$$G \xrightarrow{\eta_G} G \circ F \circ G \xrightarrow{G(\varepsilon)} G \qquad F \xrightarrow{F(\eta)} F \circ G \circ F \xrightarrow{\varepsilon_F} F$$

Then there exists a bijective $\varphi$ such that $(F, G, \varphi)$ form an adjunction between $\mathcal{C}$ and $\mathcal{D}$.

---

**Example 4.1.12.** Let $U : \textbf{R-Mod} \longrightarrow \textbf{Ab}$ be the forgetful functor, which forgets the $R$-module structure on the underlying abelian group $M$. Consider the functor $F : \textbf{Ab} \longrightarrow \textbf{R-Mod}$, where $F(A) = R \otimes A$. We'll show that this is left-adjoint to $U$ as follows.

To show this, we'll propose a morphism which we will show to be universal. If $A$ is an abelian group, then we let $\eta_A : A \longrightarrow U(F(A))$ where $\eta_A(a) = 1 \otimes a$.

Thus let $M$ be an $R$-module, and suppose there exists a morphism $f : A \longrightarrow U(M)$. Then we can define a morphism $\varphi : F(A) \longrightarrow M$ where

$$\varphi(r \otimes a) = r \cdot f(a).$$

Our construction ensures that this is a well-defined $R$-module homomorphism. Hence we clearly have the equality $U(\varphi) \circ \eta_A = f$. Visually, this becomes

$$\begin{tikzcd}
A \arrow[r, "\eta_A"] \arrow[dr, "f"'] & U(F(A)) \arrow[d, dashed, "U(\varphi)"] & F(A) \arrow[d, dashed, "\varphi"] \\
& U(M) & M
\end{tikzcd}$$

Since the construction of $\varphi$ depends directly on the existence of $f$, we see that it is unique. Hence we see that $\eta_A : A \longrightarrow U(F(A))$ is universal from $A$ to $U$. Then by Theorem 4.1, we see that we have an adjunction, so that $F$ is truly left adjoint to the forgetful functor $U$.

The following proposition is one of the main reasons why adjoint functors are extremely useful.

**Proposition 4.1.13.** Let $F, F' : \mathcal{C} \longrightarrow \mathcal{D}$ be two left adjoints of the functor $G : \mathcal{D} \longrightarrow \mathcal{C}$. Then $F, F'$ are naturally isomorphic.

*Proof.* Let $(F, G, \varphi)$ and $(F', G, \varphi')$ be two adjunctions between $\mathcal{C}$ and $\mathcal{D}$. Then these adjoints give rise to the universal morphisms

$$\eta_C : C \longrightarrow G(F(C)) \quad \eta'_C : C \longrightarrow G(F'(C))$$

for every $C \in \mathcal{C}$. Since these are both universal morphisms from $C$ to $G$, we know that they are isomorphic. Hence there exists a unique isomorphism $\theta_C : F(C) \longrightarrow F'(C)$ by universality such that $G(\theta_C) \circ \eta_C = \eta'_C$ (think of a universal diagram).

Now let $h : C \longrightarrow C'$ be a morphism in $\mathcal{C}$. Then $F'(h) \circ \theta_C = \theta_{C'} \circ F(h)$ so that the diagram

$$\begin{tikzcd}
F(C) \arrow[r, "\theta_C"] \arrow[d, "F(h)"'] & F'(C) \arrow[d, "F'(h)"] \\
F(C') \arrow[r, "\theta_{C'}"'] & F'(C')
\end{tikzcd}$$

commutes. Hence we see that $\theta : F \longrightarrow F'$ is a natural isomorphic transformation between $F$ and $F'$, so that these two functors are naturally isomorphic. ∎

The other direction holds as well. That is, two right adjoints to one left adjoint are naturally isomorphic as well, and the proof is the same. We now have our last proposition for this section.

**Proposition 4.1.14.** Let $G : \mathcal{D} \longrightarrow \mathcal{C}$ be a functor. Then $G$ has a left-adjoint $F : \mathcal{C} \longrightarrow \mathcal{D}$ if and only if for each $C \in \mathcal{C}$, the functor $\mathrm{Hom}_{\mathcal{C}}(C, G(-))$ is representable as a functor of $D \in \mathcal{D}$. Furthermore, if $\varphi : \mathrm{Hom}_{\mathcal{D}}(F_0(C), D) \cong \mathrm{Hom}_{\mathcal{C}}(C, G(D))$ is a representation of this functor, then $F_0$ is the object function of $F$.

Finally, we end this section by realizing that we can actually form composition of adjoints.

**Proposition 4.1.15.** Let $\mathcal{C}, \mathcal{D}$ and $\mathcal{E}$ be categories. Suppose we have two adjunctions as below.

$$\begin{tikzcd}
\mathcal{C} \arrow[r, shift left, "F"] & \mathcal{D} \arrow[l, shift left, "G"] \arrow[r, shift left, "F'"] & \mathcal{E} \arrow[l, shift left, "G'"]
\end{tikzcd}$$

Then the functors $F' \circ F$, $G \circ G'$ form an adjunction between $\mathcal{C}$ and $\mathcal{E}$. Further, if $(\eta, \varepsilon)$ and $(\eta', \varepsilon')$ are unit and counits of the adjunction from $(F, G)$ and $(F', G')$, then the unit and counit of the new adjunction is

$$\overline{\eta}_C = G(\eta'_{F(C)}) \circ \eta_C : C \longrightarrow (G \circ G) \circ (F' \circ F(C))$$
$$\overline{\varepsilon}_E = \varepsilon'_E \circ F'(\varepsilon_{G'(E)}) : (F' \circ F) \circ (G \circ G'(E)) \longrightarrow E$$

*Proof.* First, observe that the two given adjunctions give rise to

$$\mathrm{Hom}_{\mathcal{D}}(F(C), D) \cong \mathrm{Hom}_{\mathcal{C}}(C, G(D)) \qquad \mathrm{Hom}_{\mathcal{E}}(F'(D), E) \cong \mathrm{Hom}_{\mathcal{D}}(D, G'(E)).$$

which are relations that are natural in objects $C, D$ and $E$. Observe that in the second relation, we can set $D = F(C)$. This then translates to

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

Using the first relation, we know that $\mathrm{Hom}_{\mathcal{D}}(F(C), G(E)) \cong \mathrm{Hom}_{\mathcal{C}}(C, G(G'(E)))$. Putting this together, we then have the bijection of homsets

$$\mathrm{Hom}_{\mathcal{E}}(F' \circ F(C)), E) \cong \mathrm{Hom}_{\mathcal{C}}(C, G \circ G'(E))$$

which is natural in $C$ and $E$. Now, describing the unit and counit is a bit ugly, and not exactly necessary, since in the end we know what these adjunctions look like. The punchline here is that we can write our new unit and counit in terms of the original ones.

Observe that for any object $C$ of $\mathcal{C}$, we have the universal morphism

$$\eta_C : C \longrightarrow G(F(C)).$$

Since $F(C) \in \mathcal{D}$, we can use $\eta'$ that

$$\eta'_{F(C)} : F(C) \longrightarrow G'(F'(F(C))).$$

Finally, note that $G(\eta'_{F(C)}) : G(F(C)) \longrightarrow G(G'(F'(F(C))))$. However, we can precompose this with $\eta_C$ to have that

$$G(\eta'_{F(C)}) \circ \eta_C : C \longrightarrow G(G'(F'(F(C)))).$$

On the other hand, for any object $E$ of $\mathcal{E}$ that

$$\varepsilon'_E : F'(G'(E)) \longrightarrow E.$$

We also have $\varepsilon_D : F(G(D)) \longrightarrow D$ for any object $D \in \mathcal{D}$. Hence, we can set $D = G'(E)$ for some object $E$ of $\mathcal{E}$ to get

$$\varepsilon_{G'(E)} : F(G(G'(E))) \longrightarrow G'(E).$$

We can then get that $F'(\varepsilon_{G'(E)}) : F'(F(G(G'(E)))) \longrightarrow F'(G'(E))$. Composing this with the original $\varepsilon'_D$, we get that

$$\varepsilon'_E \circ F'(\varepsilon_{G'(E)}) : F'(F(G(G'(E)))) \longrightarrow E$$

as desired. Now showing that these remain universal is not hard. ∎

## Exercises

**1.** Give a proof of Proposition 4.1.7.

**2.** Let $U : \mathbf{Ab} \longrightarrow \mathbf{Grp}$ be the forgetful functor, and suppose $F : \mathbf{Grp} \longrightarrow \mathbf{Ab}$ is the abelianization functor. That is, if $G$ is a group and $\varphi : G \longrightarrow G'$ is a group homomorphism then

$$F(G) = G/[G,G] \qquad F(\varphi) : G/[G,G] \longrightarrow G'/[G',G'].$$

where $[G,G]$ is the commutator subgroup.

Show that we have an adjunction $\mathbf{Grp} \underset{U}{\overset{F}{\rightleftarrows}} \mathbf{Ab}$. Give a description of the unit and counits.

## 4.2 Reflective Subcategories.

**Definition 4.2.1.** Let $\mathcal{A}$ be a full subcategory of $\mathcal{C}$. We say $\mathcal{A}$ is **reflective** in $\mathcal{C}$ whenever the inclusion functor $I : \mathcal{A} \longrightarrow \mathcal{C}$ has a left adjoint $F : \mathcal{C} \longrightarrow \mathcal{A}$. We then say the functor $F$ is the **reflector**, and the adjunction $(F, I, \varphi)$ is a **reflection** of $B$.

In the case of a reflection, we obtain the bijection of hom-sets

$$\mathrm{Hom}_{\mathcal{A}}(F(C), A) \cong \mathrm{Hom}_{\mathcal{C}}(C, I(A)) \implies \mathrm{Hom}_{\mathcal{A}}(F(C), A) \cong \mathrm{Hom}_{\mathcal{C}}(C, A)$$

which is natural in both $C$ and $A$.

---

**Example 4.2.2.** Let $F : \mathbf{Grp} \longrightarrow \mathbf{Ab}$ be the abelianization functor, which sends a group $G$ to its free abelian group $G/[G, G]$. From Exercise **??**, we know that this is left adjoint to the forgetful functor $U : \mathbf{Ab} \longrightarrow \mathbf{Grp}$.

However, the functor $U : \mathbf{Ab} \longrightarrow \mathbf{Grp}$ is isomorphic to the inclusion functor $I : \mathbf{Ab} \longrightarrow \mathbf{Grp}$. Hence, $F$ is also left adjoint to the inclusion functor, so that $\mathbf{Ab}$ is a reflective subcategory of $\mathbf{Grp}$.

---

---

**Example 4.2.3.** Let $\mathbf{Top}$ be the category of topological spaces with morphisms continuous functions. Let $\mathbf{CHaus}$, the category of compact Hausdorff spaces, which is a subcategory of $\mathbf{Top}$.

If we let $X$ be a topological space, then we denote $\beta(X)$ to be the Stone-Cech compactification. Let $I : \mathbf{CHaus} \longrightarrow \mathbf{Top}$ be the inclusion functor. Then the definition of the Stone-Cech compactification of a space $X$ is the universal property:

$$\begin{tikzcd}
X \arrow[r, "u"] \arrow[dr, "f'"'] & I(\beta(X)) \arrow[d, dashed, "\beta(f)"] & \beta(X) \arrow[d, dashed, "f"] \\
& I(C) & C
\end{tikzcd}$$

That is, the Stone-Cech compactification is a topological space $\beta(X)$ with a morphism $u : X \longrightarrow \beta(X)$ which is universal across all morphisms $f : X \longrightarrow C$ where $C$ is compact, Hausdorff.

Thus we see that a Stone-Cech compactification gives rise to an object $\beta(X) \in \mathbf{CHaus}$ and a universal morphism $X \longrightarrow I(\beta(X))$ from $X$ to $I$. Now by Proposition 4.1, this makes $\beta : \mathbf{Top} \longrightarrow \mathbf{CHaus}$ a functor, which is left adjoint to the inclusion functor $I : \mathbf{CHaus} \longrightarrow \mathbf{Top}$.

This then makes $\beta : \mathbf{Top} \longrightarrow \mathbf{CHaus}$ a reflector, so that the adjunction is a reflection between $\mathbf{Top}$ and $\mathbf{CHaus}$. Consequently we have the bijection

$$\mathrm{Hom}_{\mathbf{Top}}(X, I(C)) \cong \mathrm{Hom}_{\mathbf{CHaus}}(\beta(X), C) \implies \mathrm{Hom}_{\mathbf{Top}}(X, C) \cong \mathrm{Hom}_{\mathbf{CHaus}}(\beta(X), C).$$

since $I(C)$ is technically no different than from $C$. This bijection is natural in both $X$ and $C$.

---

---

**Example 4.2.4.** Let $\mathbf{Ab_{TF}}$ represent the category of abelian groups with torsion free elements (for a lack of better notation). Then we have a natural inclusion functor $I : \mathbf{Ab_{TF}} \longrightarrow \mathbf{Ab}$. Now consider the functor $F : \mathbf{Ab} \longrightarrow \mathbf{Ab_{TF}}$, which we define as follows:

**Objects.** Let $G$ be an abelian group. Then $F(G) = G_{TF}$ where

$$G_{TF} = \{g \in G \mid g^n \neq e \text{ for } n = 1, 2, 3, \dots\}.$$

That is, it sends $G$ to its underlying abelian group of torsion-free elements. It's not hard to show this is an abelian group.

**Morphisms.** Suppose $\varphi : G \longrightarrow H$ is a morphism between abelian groups. Then we set $F(\varphi) = \varphi_{TF}$ where

$$\varphi_{TF} : G_{TF} \longrightarrow H_{TF} \qquad \varphi_{TF}(g) = \varphi(g).$$

Note that this definition will cause no issues, since $\text{ord}(g) = \text{ord}(\varphi(g))$. Thus we simply obtain $\varphi_{TF}$ by restricting $\varphi$ to $G_{TF}$.

To show that $F$ is left adjoint to $I$, we need to demonstrate that there exists a universal morphism $\eta_G : G \longrightarrow I(F(G))$ for every $G \in \mathbf{Ab}$. Hence we propose $\eta_G$ takes on the form

$$\eta_G(g) = \begin{cases} g & \text{if } \text{ord}(g) = \infty \\ e & \text{otherwise.} \end{cases}$$

To show this is universal from $G$ to $I$, suppose we have a morphism $\varphi : G \longrightarrow I(H)$, where $H \in \mathbf{Ab_{TF}}$. Then there exists a morphism $\psi : F(G) \longrightarrow H$ such that $I(\varphi) \circ \eta_G = \varphi$. Visually, that is,

$$\begin{tikzcd} G \arrow[r, "\eta_G"] \arrow[dr, "\varphi"'] & I(F(G)) \arrow[d, dashed, "I(\psi)"] \\ & I(H) \end{tikzcd} \qquad \begin{tikzcd} F(G) \arrow[d, dashed, "\psi"] \\ H \end{tikzcd}$$

Sure such a morphism exists, but why the equality?

***$g \in$ Ker $(\eta_G)$*.** If $g \in \text{Ker}(\eta_G)$, then $g$ has finite order. Hence we see that $\varphi(g) = e$; this is because $\text{ord}(\varphi(g)) = \text{ord}(g) < \infty$, but the only element in $I(H)$ with finite order is $e$. We then have that $g \in \text{Ker}(\varphi)$. Therefore,

$$I(\psi) \circ \eta_G(g) = I(\psi)(e) = e = \varphi(g).$$

Hence $I(\psi) \circ \eta_G = \varphi$ if $g \in \text{Ker}(\eta_G)$.

***g* $\notin$ Ker ($\eta_G$).** if $g \notin \text{Ker}(\eta_G)$, then we know that $\text{ord}(g) = \infty$. Therefore, we see that

$$I(\psi) \circ \eta_G(g) = I(\varphi)(g) = \varphi(g).$$

Hence $I(\psi) \circ \eta_G = \varphi$ for $g \notin \text{Ker}(\eta_G)$.

By our previous work, we then have that $I(\psi) \circ \eta_G = \varphi$, as desired. Now $\psi$ is of course unique based on its construction, since its definition depends directly on $\varphi$. We then have that $\eta_G : G \longrightarrow I(F(G))$ is universal from $G$ to $I$ for each $G \in \mathbf{Ab}$!

We then have by Theorem 4.1 that $F, I$ form an adjunction, so that $F$ is the left adjoint of $I$. Hence by definition, we see that $\mathbf{AB_{TF}}$ forms a full reflective subcategory of $\mathbf{Ab}$.

---

## Exercises

**1.** Is **FinSet** a reflective subcategory of **Set**?

**2.** Let $G$ and $H$ be a groups. Prove that

$$G * H/[G * H, G * H] \cong G/[G, G] \oplus H/[H, H]$$

where $G * H$ denotes the free product of $G$ and $H$. (What this is saying is that $F : \mathbf{Grp} \longrightarrow \mathbf{Ab}$, the abelianization functor, preserves coproducts. Eventually, this fact will immediately follow by our knowledge of the adjunction $\mathbf{Grp} \underset{U}{\overset{F}{\rightleftarrows}} \mathbf{Ab}$. )

## 4.3 Equivalence of Categories

In an ideal world, if we have a category of which we are interested in, our goal would be to find an isomorphism between it and a category of which we understand very well. We then know that certain mathematical structures are invariant between transitioning between the two, so that we could better understand our desired category.

However, this is generally too much to ask for. Many categories which are constructed are constructed in such a way that they're not isomorphic to anything we're familiar with; if they were, then they probably wouldn't be interesting. Hence we have a more useful notion of equivalence between categories.

**Definition 4.3.1.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor. We say that $\mathcal{C}$ is **equivalent** to $\mathcal{D}$ if there exists a functor $G : \mathcal{D} \longrightarrow \mathcal{C}$ and natural isomorphisms $\eta : I_{\mathcal{C}} \longrightarrow G \circ F$ and $\varepsilon : F \circ G \longrightarrow I_{\mathcal{D}}$.

In this case, we say both $F$ and $G$ are an **equivalence of categories**.

---

**Example 4.3.2.** Let $X$ and $Y$ be sets, and regard them as discrete categories. Then a functor $F : X \longrightarrow Y$ is just a function between sets. In this case, to say that $X$ and $Y$ are equivalent is if there exists a functor (function!) $G : Y \longrightarrow X$ such that we have natural isomorphisms $\eta_x : x \longrightarrow G(F(x))$ and $\varepsilon_x : F(G(x)) \longrightarrow x$. However, each category has nontrivial morphisms; hence we see that each of these must be identity morphisms so that

$$G(F(x)) = x \qquad F(G(x)) = x.$$

What this then means is that an equivalence of categories for sets is just a pair of invertible functions. That is, it gives rise to an isomorphism.

---

Since $\eta, \varepsilon$ are already natural transformations, this simply makes them natural isomorphisms. It turns out that the notion of equivalence is more useful than of an isomorphism. An isomorphism is just too much to ask, but equivalence does give us nice invariants too.

**Definition 4.3.3.** A **adjoint equivalence** between categories $C$ and $D$ is an adjunction $(F, G, \eta, \varepsilon)$ where the unit and counit $\eta$ and $\varepsilon$ are natural isomorphisms.

It turns our an adjoint equivalence is the same thing as an equivalence between categories. But before we move on, we prove a lemma and a proposition.

**Lemma 4.3.4.** Let $\mathcal{C}$ be a category, and $f : A \longrightarrow B$ a morphism. Then $f$ induces a natural transformation

$$f^* : \mathrm{Hom}_{\mathcal{C}}(C, -) \longrightarrow \mathrm{Hom}_{\mathcal{C}}(C', -)$$

Then $f^*$ is a monomorphism if and only if $f$ is an epimorphism, and $f^*$ is an epimorphism if and only if $f$ is a split monomorphism (that is, if and only if $f$ has a left-inverse.)

*Proof.*

$\Longrightarrow$ Observe that $\text{Hom}_{\mathcal{C}}(C, -) \longrightarrow \mathcal{C} \longrightarrow \textbf{Set}$ is a functor. Then $f^* : \text{Hom}_{\mathcal{C}}(C, -) \longrightarrow \text{Hom}_{\mathcal{C}}(C', -)$ is a natural transformation where $f : C' \longrightarrow C$. Now suppose $\eta, \eta' : F \longrightarrow \text{Hom}_{\mathcal{C}}(C, -)$, where $F : \mathcal{C} \longrightarrow \textbf{Set}$ is a functor, are natural transformations. Then if $f^*$ is monic,

$$f^* \circ \eta = f^* \circ \eta' \implies \eta = \eta'.$$

Now let $h : A \longrightarrow A'$ be a morphism in $\mathcal{C}$. Then we have the commutative diagram

$$\begin{tikzcd}
A \arrow[d, "h"] & F(A) \arrow[r, "\eta_A, \eta'_A"] \arrow[d, "F(h)"'] & \text{Hom}_{\mathcal{C}}(C, A) \arrow[r, "f^*"] \arrow[d, "h_*"] & \text{Hom}_{\mathcal{C}}(C', A) \arrow[d, "h_*"] \\
A' & F(A') \arrow[r, "\eta_{A'}, \eta'_{A'}"'] & \text{Hom}_{\mathcal{C}}(C, A') \arrow[r, "f^*"'] & \text{Hom}_{\mathcal{C}}(C', A')
\end{tikzcd}$$

where we denote $\eta_A, \eta'_A$ on the arrow to signify the fact that both $\eta_A, \eta'_A$ are morphisms from $F(A)$ to $\text{Hom}_{\mathcal{C}}(C, A)$. Now let $x \in F(A)$. Then

$$f^* \circ \eta_A(x) = f^* \circ \eta'_A(x) \iff \eta_A(x) \circ f = \eta'_A(x) \circ f.$$

But if $f$ is monic, then $f^* \circ \eta_A(x) = f^* \circ'_A (x)$ implies that $\eta_A = \eta'_A$. Hence we see that $\eta_A(x) \circ f = \eta'_A(x) \circ f \implies \eta_A(x) = \eta'_A(x)$.

$\Longleftarrow$ Now suppose $f$ is epic. Then using the same notation as earlier, note that

$$f^* \circ \eta_A(x) = f^* \circ \eta'_A(x) \iff \eta_A(x) \circ f = \eta'_A(x) \circ f \implies \eta_A = \eta_A.$$

Hence we see that $f^*$ is a monomorphism.

∎

Taking the dual of what we proved, we prove the second part of the lemma. Now we'll use this lemma in the theorem below, one which will be very useful.

**Proposition 4.3.5.** Let $(F, G, \eta, \varepsilon)$ be an adjunction between categories $\mathcal{C}$ and $\mathcal{D}$. Then
***(i)*** $G$ is faithful if and only if for each $D \in \mathcal{D}$, $\varepsilon_D$ is epic
***(ii)*** $G$ is full if and only if every $\varepsilon_D$ is split monic.
Therefore, $G$ is full and faithful if and only if $\varepsilon_D$ is an isomorphism between $F(G(D))$ and $D$.

*Proof.* If $G : \mathcal{D} \longrightarrow \mathcal{C}$ is a functor, then we see that $G$ itself becomes a natural transformation between the two functors:

$$G_{D,-} : \text{Hom}_{\mathcal{D}}(D, -) \longrightarrow \text{Hom}_{\mathcal{D}}(G(D), G(-)).$$

Recall that we have an adjunction given by $F, G$. Then there exists a bijection $\varphi$ where

$$\varphi_{C,D'} : \text{Hom}_{\mathcal{C}}(F(C), D') \longrightarrow \text{Hom}_{\mathcal{D}}(C, G(D)).$$

Thus $\varphi^{-1} : \mathrm{Hom}_{\mathcal{D}}(C, G(D)) \longrightarrow \mathrm{Hom}_{\mathcal{D}}(F(C), D')$. Moreover, if $D$ is an arbitrary object, this becomes a natural transformation between the two functors:

$$\varphi^{-1}_{C,-} : \mathrm{Hom}_{\mathcal{D}}(C, G(-)) \longrightarrow \mathrm{Hom}_{\mathcal{C}}(F(C), -).$$

Let $C = G(D)$. Then we have the following sequence of natural transformations:

$$\mathrm{Hom}_{\mathcal{D}}(D, -) \xrightarrow{G_{D,-}} \mathrm{Hom}_{\mathcal{C}}(G(D), G(-)) \xrightarrow{\varphi^{-1}_{G(D),G(-)}} \mathrm{Hom}_{\mathcal{D}}(F(G(D)), -)$$

Composing the natural transformations, we finally obtain a natural transformation $\varphi^{-1}_{G(D),G(-)} \circ G_{D,-} : \mathrm{Hom}_{\mathcal{D}}(D, -) \longrightarrow \mathrm{Hom}_{\mathcal{D}}(F(G(D)), -)$. How is this natural transformation given? We can assign $-$ as $D$ itself, and see what happens when we consider the identity morphism $1_D : D \longrightarrow D$. In this case

$$\varphi^{-1}_{G(D),G(D)} \circ G_{D,D}(1_D) = \varphi^{-1}_{G(D),G(D)}(1_{G(D)}) = \varepsilon_D$$

by definition of the counit $\varepsilon_D$. Now we understand how this poorly-notated natural transformation works! In general, for and $f : D \longrightarrow D'$, we see that

$$\varphi^{-1}_{G(D),G(D')} \circ G_{D,D'}(f) = f \circ \varepsilon_D. \tag{4.1}$$

Thus, we see that this natural transformation is in disguise; it's actually just $\varepsilon^*_D$ : $\mathrm{Hom}_{\mathcal{D}}(D, -) \longrightarrow \mathrm{Hom}_{\mathcal{D}}(F(G(D), -)$!

***(i)*** $\iff$ If $G$ is faithful, then the natural transformation in equation (7) is one to one. This makes $\varepsilon^*_D$ a monomorphism. By the previous lemma, this holds if and only if $\varepsilon_D$ is epic for every $D$ in $\mathcal{D}$.

***(ii)*** $\iff$ On the other hand, if $G$ is full, then this natural transformation in equation (7) surjective. This makes $\varepsilon^*_D$ an epimorphism, and by the previous lemma, that holds if and only if $\varepsilon_D$ is a split monomorphism.

∎

**Theorem 4.3.6.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor. Then the following are equivalent.

***(i)*** $G$ is an equivalence of categories

***(ii)*** $G$ is part of an adjunction $(F, G, \eta, \varepsilon)$ where $\eta, \varepsilon$ are natural isomorphisms

***(iii)*** $F$ is full and faithful, and each object $C$ is isomorphic to $G(D)$ for some object $D$.

Note that this theorem is symmetric; one could interchange $G$ with $F$, and then obtain the same exact results. Thus, one way of stating this theorem is that $\mathcal{C}$ and $\mathcal{D}$ are equivalent as categories if and only if there exits full and faithful functors $F : \mathcal{C} \longrightarrow \mathcal{D}$ and $G : \mathcal{D} \longrightarrow \mathcal{C}$; or if and only if $F, G$ form an adjoint equivalence.

*Proof.*
***(i)*** $\implies$ ***(iii)*** Suppose we have an equivalence of categories given by $F : \mathcal{C} \longrightarrow \mathcal{D}$ and $G : \mathcal{D} \longrightarrow \mathcal{C}$, with natural isomorphisms

$$\varphi : F \circ G \cong I_{\mathcal{D}} \qquad \psi : G \circ F \cong I_{\mathcal{C}}.$$

Let $f : C \longrightarrow C'$ be a morphism in $\mathcal{C}$. Then observe that the following diagram

$$\begin{tikzcd} C \arrow[d, "f"] \\ C' \end{tikzcd} \qquad \begin{tikzcd} G(F(C)) \arrow[r, "\psi_C"] \arrow[d, "G(F(C))"'] & C \arrow[d, "f"] \\ G(F(C')) \arrow[r, "\psi'_C"'] & C' \end{tikzcd}$$

is commutative. In an equations, we have that $f = \psi'_C \circ G(F(f)) \circ \psi_{C'}^{-1}$. Thus suppose that $f_1, f_2 : C \longrightarrow C'$ are two morphisms such that $F(f_1) = F(f_2)$. Then we get a pair of commutative diagrams, similar to the ones above, which translate into the equations

$$f_1 = \psi'_C \circ G(F(f_1)) \circ \psi_{C'}^{-1} \qquad f_2 = \psi_{C'} \circ G(F(f_2)) \circ \psi_{C'}^{-1}.$$

Then if $F(f_1) = F(f_2)$, the above equations guarantee that $f_1 = f_2$. Hence we see that $F$ is a faithful functor. Since the statement is symmetric in both $F$ and $G$, we have also that $G$ is faithful.
To show that $F$ is full, suppose there exists a morphism $h : F(C) \longrightarrow F(C')$ for a pair of objects $C, C'$. Let $f = \psi_{C'} \circ G(h) \circ \psi_C$. Then we have the commutative squares

$$\begin{tikzcd} G(F(C)) \arrow[r, "\psi_C"] \arrow[d, "G(h)"'] & C \arrow[d, "f"] \\ G(F(C')) \arrow[r, "\psi'_C"'] & C'. \end{tikzcd} \qquad \begin{tikzcd} G(F(C)) \arrow[r, "\psi_C"] \arrow[d, "G(F(f))"'] & C \arrow[d, "f"] \\ G(F(C')) \arrow[r, "\psi'_C"'] & C'. \end{tikzcd}$$

and hence we have that $G(h) = G(F(f))$. But since $G$ is faithful, this implies that $h = F(f)$. Hence we have that there exists a $f' : C \longrightarrow C'$ such that $h = F(f)$, so that $F$ is full. Again, by symmetry, we have that $G$ is full, as desired.
Now since $\varphi : G \circ F \cong I_{\mathcal{C}}$, we see that every object $C$ is assigned an isomorphism $\varphi_C : G(F(C)) \longrightarrow C$. Hence every object $C$ is isomorphic to some $G(D)$ where $D = F(C)$. Similarly, since $\psi : F \circ G \cong I_{\mathcal{D}}$, we know that each object $D$ is assigned an isomorphism $\psi_D : F(G(D)) \longrightarrow D$. Hence every object $D$ is isomorphic to some object $F(C)$ for $C = G(D)$.
***(iii)*** $\implies$ ***(ii)*** Suppose $(iii)$ holds. For any arbitrary object $C \in \mathcal{C}$, there exists an isomorphism $\eta_C : C \longrightarrow G(D)$ for some object $D \in \mathcal{D}$. Denote such an object as $F_0(C)$. Now consider any other morphism $g : C \longrightarrow G(D')$. Then we have that

$$
\begin{tikzcd}
C \arrow[r, "\eta_C"] \arrow[dr, "g"'] & G(F_0(C)) \arrow[d, dashed, "g\circ f^{-1}"] \\
& G(D')
\end{tikzcd}
$$

is commutative. Now since $g \circ \eta_C^{-1} : G(F_0(C)) \longrightarrow G(D')$, and because $G$ is full, we know that there exists a $h : F_0(C) \longrightarrow D'$ such that $g \circ \eta_C^{-1} = G(h)$. To show that this is unique, suppose there existed another $k : G(F_0(C)) \longrightarrow G(D')$ such that $g = k \circ \eta_C$. Then by the same argument, there exists a $h' : F_0(C) \longrightarrow D'$ such that $G(h') = k$. Furthermore, we'll have that

$$k = G(h') = g \circ \eta_C^{-1} \qquad G(h) = g \circ \eta_C^{-1}$$

so that $G(h') = G(h)$. However, since $G$ is faithful, we have that $h' = h$. Hence, $h$ is unique!

Since $h$ is unique, this implies that $\eta_C : C \longrightarrow G(F_0(C))$ is universal from $C$ to $G$. Since such a universal isomorphism exists for each object of $C$, we have by Proposition 4.1 that there exists a functor $F : \mathcal{C} \longrightarrow \mathcal{D}$ with object function $F_0(C)$ which is left adjoint to $G$. Hence we have an adjunction $(F, G, \eta', \varepsilon)$. However, since universal morphisms are unique, we see that $\eta' = \eta$, so that $\eta$, our unit, is a natural isomorphism.

Finally, observe that for any object $D$, we have that

$$G(\varepsilon_D) \circ \eta_{G(D)} = 1_{G(D)}$$

for our adjunction. Since $\eta_{G(D)}$ is an isomorphism, we have that $G(\varepsilon_D) = \eta_{G(D)}^{-1}$. Sine $G$ is full and faithful, we see that $\varepsilon_D$ must be an isomorphism as well.

Thus, in total, we have an adjoint equivalence $(F, G, \eta, \varepsilon)$, as desired.

***(ii)* $\implies$ *(i)*** This direction is clear, since an adjoint equivalence automatically establishes an equivalence of categories.

With $(i) \implies (iii) \implies (ii) \implies (i)$, we see that all of the conditions are equivalent. ∎

---

**Example 4.3.7.** Let $R$ and $S$ be rings and consider the categories $R$-**Mod** and $S$-**Mod**. Then there are two different "product" categories we can form: The categories $(R \times S)$-**Mod** and $R$-**Mod** $\times$ $S$-**Mod**

---

Next, we introduce some properties of equivalences.

**Proposition 4.3.8.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be an equivalence of categories with the corresponding inverse functor $G : \mathcal{D} \longrightarrow \mathcal{C}$. Let $f : C \longrightarrow C'$ be a morphism in $\mathcal{C}$. Then

*(i)* $f$ is a monomorphism (epimorphism) if and only if $F(f)$ is a monomorphism (epimorphism)

*(ii)* $C$ is initial (terminal) if and only if $F(C)$ is initial (terminal).

Consequently, we have that $f$ is an isomorphism (a monomorphism and epimorphism) if and only if $F(F)$ is an isomorphism. Note this is not generally true! Additionally, we also have that $C$ is a zero object (terminal and initial) if and only if $F(C)$ is a zero object. Finally, observe that this proposition is symmetric, so that the same conclusions hold for morphisms and objects in $\mathcal{D}$ governed by $G : \mathcal{D} \longrightarrow \mathcal{C}$.

*Proof.*

***(i)*** $\implies$ Suppose $f : C \longrightarrow C'$ is a monomorphism. Consider two morphisms $g, h : D \longrightarrow F(C)$ such that $F(f) \circ g = F(f) \circ h$. By the previous theorem, we know however that there exists an object $A$ of $\mathcal{C}$ such that $D \cong F(A)$. Hence there exists an isomorphism $\theta : F(A) \longrightarrow D$. We then have the diagram:

$$F(A) \xrightarrow{\theta} D \underset{g}{\overset{h}{\rightrightarrows}} F(C) \xrightarrow{F(f)} F(C')$$

Note that $h \circ \theta, g \circ \theta : F(A) \longrightarrow F(C)$. Since $F$ is full, we know that there exists morphism $k, k' : A \longrightarrow C$ such that $g \circ \theta = F(k)$ and $h \circ \theta = F(k')$. Now observe that

$$F(f \circ k) = F(f) \circ F(k) = F(f) \circ h \circ \theta$$
$$F(f \circ k') = F(f) \circ F(k') = F(f) \circ g \circ \theta.$$

However, since $F(f) \circ h = F(f) \circ g$, we see that $F(f \circ k) = F(f \circ k')$. However, since $F$ is faithful, we have that $f \circ k = f \circ k'$. But since $f$ is a monomorphism, we have that $k = k'$. Hence $F(k) = F(k') \implies g \circ \theta = k \circ \theta$, and since $\theta$ is an isomorphism, we have that $h = g$. Therefore, $F(f)$ is also monic.

$\impliedby$ Suppose $f : C \longrightarrow C'$ and $F(f)$ is monic. Consider two morphism $g, h : A \longrightarrow C'$ in $\mathcal{C}$, and suppose that $f \circ g = f \circ k$. Then $F(f) \circ F(g) = F(f) \circ F(k) \implies F(g) = F(k)$, since $F(f)$ is monic. However, $F$ is faithful, so that $g = k$. Hence $f$ is monic as well.

***(ii)*** $\implies$ Suppose $C$ is initial in $\mathcal{C}$. Let $D$ be an object in $\mathcal{D}$. Then observe that, since $\mathcal{C}$ and $\mathcal{D}$ are equivalent, there exists an isomorphism $\theta : F(A) \longrightarrow D$ for some object $A$ of $\mathcal{C}$. Since $C$ is initial, we know that there exists a unique morphism $f_C : C \longrightarrow A$. Hence $F(f_C) : F(C) \longrightarrow F(A)$. We then have that $F(f_c) \circ \theta : F(C) \longrightarrow D$. Hence there exists a morphism from $F(C)$ to $D$.

Now suppose $f_1, f_2; F(C) \longrightarrow D$. Then $\theta^{-1} \circ f_1, \theta^{-1} \circ f_2 : F(C) \longrightarrow F(A)$. Since $F$ is full, we know that there exist morphism $k_1, k_1 : C \longrightarrow A$ such that $F(k_1) = \theta^{-1} \circ f_1$ and $F(k_2) = \theta^{-1} \circ f_2$. However, since $C$ is initial, we see that $k_1 = k_2 = f_C$. Hence $f_1 = f_2$, so that there is exactly one morphism $f_1 = f_2 : F(C) \longrightarrow D$.

Since $D$ was an arbitrary object of $\mathcal{D}$, we have that $F(C)$ is initial.

$\impliedby$ Suppose $F(C)$ is an initial object. Consider any object $C'$ of $\mathcal{C}$. Then since $F(C)$ is initial, there exists a unique morphism $f : F(C) \longrightarrow F(C')$. Since $F$ is full, we know that this corresponds with a morphism $k : C \longrightarrow C'$ such that $F(k) = f$.

Hence we have a unique morphism $k : C \longrightarrow C'$. And since $C'$ was an arbitrary object of $\mathcal{C}$, we have that $C$ is initial, as desired.

∎

The proofs in which we proved $f$ to be an epimorphism, and for $C$ to be a terminal object, are very similar. This proposition will soon be generalized, but this gives us insight into how useful the concept of equivalent categories truly is.

## 4.4 Adjoints on Preorders.

Interesting things happen when one applies adjoint concepts to functors between preorders; ones which preserve order in a special way. It's actually often the case where we have two mathematical structures involving chains of arrows which reverse when transferring between one and the other. We give such a concept a definition first, before introducing a theorem about such structures.

**Definition 4.4.1.** Let $\mathcal{P}$ and $\mathcal{Q}$ be two preorders. If there exists functors $F : \mathcal{P} \longrightarrow \mathcal{Q}$ and $G : \mathcal{Q} \longrightarrow \mathcal{P}$ such that

$$F(P) \leq Q \iff P \leq G(Q),$$

That is, there exists $f : F(P) \longrightarrow Q$ if and only if there exists $g : P \longrightarrow G(Q)$, then $F$ and $G$ are called a **monotone Galois connection**. On the other hand, if we have that

$$F(P) \leq Q \iff P \geq G(Q)$$

then $F$ and $G$ are called a **antitone Galois connection**.

**Theorem 4.4.2.** Let $\mathcal{P}, \mathcal{Q}$ be two preorders, and suppose $F : \mathcal{P} \longrightarrow \mathcal{Q}^{\text{op}}$ and $G : \mathcal{Q}^{\text{op}} \longrightarrow \mathcal{P}$ are two order preserving functors. Then $F$ is left adjoint to $G$ if and only if for all $P \in \mathcal{P}$ and $Q \in \mathcal{Q}$

$$F(P) \geq Q \iff P \leq G(Q).$$

Given such an adjunction, we then have that our unit establishes $P \leq G(F(P))$ and the counit establishes $F(G(Q)) \leq Q$.

*Proof.* Observe that if $F$ is left adjoint to $G$, then we have the bijection

$$\text{Hom}_{\mathcal{Q}^{\text{op}}}(F(P), Q) \cong \text{Hom}_{\mathcal{P}}(P, G(Q))$$

which gives rise to the desired correspondence; on the other hand, such a bijection gives rise to an adjunction. With such an adjunction, we know that for each $P, Q$, there exist morphisms $\eta_P : P \longrightarrow G(F(P))$ and $\varepsilon_Q : F(G(Q)) \longrightarrow Q$. Hence $P \leq G(F(P))$ and $F(G(Q)) \geq Q$. ∎

The above theorem came out of the observation that there is a connection between fields, their subfields, and their groups of automorphisms, an observation which arises in Galois Theory. The goal of Galois Theory is to understand polynomials and their roots; when they can be factorized, when and where we can find their roots. The study of Galois groups is now used widely in number theory. For example, part of Andrew Wiles' work in proving Fermat's Last Theorem involved Galois representations.

It was this theorem, rooted in Galois Theory, that motivated the Theorem 4.**??** at the beginning of this section. The Fundamental Theorem of Galois Theory is simply a *stronger*, special case, since in this case, the functors are literally inverses of each other. The theorem we introduced, however, simply requires the functors to be adjoints of one another.

---

**Example 4.4.3.** Let $U, V$ be sets, and observe that their power sets $\mathcal{P}(U)$ and $\mathcal{P}(V)$ form categories; specifically, preorders, ordered by set inclusion.

Suppose $f : U \longrightarrow V$ is a function in **Set**. Then $f$ induces a functor $f_* : \mathcal{P}(U) \longrightarrow \mathcal{P}(V)$, where

$$f_*(X) = \{f(x) \mid x \in X\}.$$

Note that if $X \subseteq X'$, then $f_*(X) \subseteq f_*(X')$. Hence this is an order-preserving functor. Now observe that $f$ also induces a functor $f^* : \mathcal{P}(V) \longrightarrow \mathcal{P}(U)$ where

$$f^*(Y) = \{x \mid f(x) \in Y\}.$$

Note that this also preserves order. In addition, we have that if $f_*(X) \leq Y$, then this holds if and only if $f(X) \subseteq Y$. We then have that this holds if and only if $X \subseteq f_*(Y)$, Hence we have a Galois connection, so that we may apply Theorem 4.**??** to conclude that $f_*$ is left adjoint to $f^*$.

---

# 4.5 Exponential Objects and Cartesian Closed Categories.

Before we introduce the notion of cartesian closed category, we begin with a preliminary proposition.

**Proposition 4.5.1.** Suppose $\mathcal{C}$ is a category, and consider the functors

$$U : C \longrightarrow \mathbf{1} \qquad \Delta : \mathcal{C} \longrightarrow \mathcal{C} \times \mathcal{C}.$$

where $\mathbf{1}$ is the one object category.
***(i)*** If $U$ has a left adjoint, then $\mathcal{C}$ has an initial object.
***(ii)*** If $\Delta$ has a left adjoint, then $\mathcal{C}$ has finite coproducts.
***(iii)*** If $U$ has a right adjoint, then $\mathcal{C}$ has a terminal object.
***(iv)*** If $\Delta$ has a right adjoint, then $\mathcal{C}$ has finite products.

The proof is a straightforward, although tedious, so we sketch it out as follows.

*Proof.*
**Adjoints of $U$.** First, let $F : \mathbf{1} \longrightarrow \mathcal{C}$ be a left adjoint of $U$. Suppose $F(1) = I$ in $\mathcal{C}$. Then for any $C \in \mathcal{C}$, we have the bijection $\text{Hom}_{\mathcal{C}}(F(1), C) \cong \text{Hom}_{\mathbf{1}}(1, U(C))$ which implies that

$$\text{Hom}_{\mathcal{C}}(I, C) \cong \text{Hom}_{\mathbf{1}}(1, 1).$$

In other words, for each object $C$, there is exactly one and only one morphism $i_C : I \longrightarrow C$, which makes $I$ an initial object.

On the other hand, suppose $G : 1 \longrightarrow \mathcal{C}$ is a right adjoint of $U$. Then if $G(1) = T$, we have the bijection $\text{Hom}_{\mathbf{1}}(U(C), 1) \cong \text{Hom}_{\mathcal{C}}(C, G(1))$ which implies that

$$\text{Hom}_{\mathbf{1}}(1, 1) \cong \text{Hom}_{\mathcal{C}}(C, T)$$

so that for each object $C$ there exists a unique morphism $t_C : C \longrightarrow T$, which makes $T$ a terminal object. Hence left and right adjoints guarantee the existence of initial and terminal objects.

**Adjoints of $\Delta$.** Let $F : \mathcal{C} \times \mathcal{C} \longrightarrow \mathcal{C}$ be a left adjoint of $\Delta$, so that we have the relation

$$\begin{tikzcd} \mathcal{C} \times \mathcal{C} \arrow[r, shift left, "F"] & \mathcal{C} \arrow[l, shift left, "\Delta"] \end{tikzcd}$$

Then for each object $(A, B) \in \mathcal{C} \times \mathcal{C}$, we have the morphism $\eta_{(A,B)} : (A, B) \longrightarrow \Delta(F(A, B))$, which we can rewrite as $\eta_{(A,B)} : (A, B) \longrightarrow (F(A, B), F(A, B))$. We can put this into a universal diagram

$$\begin{tikzcd} (A, B) \arrow[r, "\eta_{(A,B)}"] \arrow[dr, "f"'] & (F(A, B), F(A, B)) \arrow[d, dashed, "(g', g')"] \\ & (C, D) \end{tikzcd} \qquad \begin{tikzcd} (A, B) \arrow[r, "(i,j)"] \arrow[dr, "f"'] & (A \amalg B, A \amalg B) \arrow[d, dashed, "(f', f')"] \\ & (C, D) \end{tikzcd}$$

where the diagram on the right is the coproduct diagram of $A \times B$. Since both of the pairs $\Big((F(A,B), F(A,B)), \eta_{(A,B)}\Big)$ and $\Big((A \times B, A \times B), (\pi_A, \pi_B)\Big)$ are universal from $(A,B)$ to $\Delta$, they must be isomorphic. As two universal objects are isomorphic, we therefore have,

$$F(A,B) \cong A \amalg B$$

so that a left adjoint gives rise to products.

Let $G : \mathcal{C} \times \mathcal{C} \longrightarrow \mathcal{C}$ be a right adjoint of $\Delta$, so that we have

$$\begin{tikzcd} \mathcal{C} \arrow[r, shift left, "\Delta"] & \mathcal{C} \times \mathcal{C} \arrow[l, shift left, "G"] \end{tikzcd}$$

The adjunction gives rise to a universal morphism $\varepsilon_{(A,B)} : \Delta(G(A,B)) \longrightarrow (A,B)$, which we can rewrite as $\varepsilon_{(A,B)} : (G(A,B), G(A,B)) \longrightarrow (A,B)$. We then have the diagram

$$\begin{tikzcd} (A,B) & (G(A,B), G(A,B)) \arrow[l, "\varepsilon_{(A,B)}"'] \\ & (C,D) \arrow[ul, "f"] \arrow[u, dashed, "(g',g')"'] \end{tikzcd} \qquad \begin{tikzcd} (A,B) & (A \times B, A \times B) \arrow[l, "(\pi_A,\pi_B)"'] \\ & (C,D) \arrow[ul, "f"] \arrow[u, dashed, "(f',f')"'] \end{tikzcd}$$

where the diagram on the right is the product diagram of $A \times B$. Thus we see that $\Big((G(A,B), G(A,B)), \varepsilon_{(A,B)}\Big)$ and $\Big((A \times B, A \times B), (\pi_A, \pi_B)\Big)$ are both universal from $\Delta$ to $(A,B)$. As universal objects from the same construction are isomorphic, we have that

$$G(A,B) \cong A \times B$$

so that this adjunction gives rise to coproducts. ∎

Thus if we have left and right adjoints of the functors $U$ and $\Delta$, we get initial and terminal objects as well as finite products and coproducts. Note, however, that finite products require (and give rise to) initial objects, and similarly that finite coproducts require (and give rise to) terminal objects.

Next, we make the following definition.

**Definition 4.5.2.** Let $\mathcal{C}$ be a category with finite products. Suppose $Y, Z$ are objects in $\mathcal{C}$. We say $Z^Y$ is an **exponent object** in $\mathcal{C}$ if there exists a morphism $\textbf{eval} : (Z^Y \times Y) \longrightarrow Z$ which is universal from $- \times Y : \mathcal{C} \longrightarrow \mathcal{C}$ to the object $Z$.

Visually, this translates into requiring that the following diagram commutes.

$$\begin{tikzcd} Z & Z^Y \times Y \arrow[l, "\textbf{eval}"'] \\ & X \times Y \arrow[ul, "g"] \arrow[u, dashed, "(h, \mathrm{id}_Y)"'] \end{tikzcd} \qquad \begin{tikzcd} Z^Y \\ X \arrow[u, dashed, "h"'] \end{tikzcd}$$

Hence, every morphism, with the domain being any product with $\boldsymbol{Y}$, and codomain being $\boldsymbol{Z}$, uniquely factors through $\boldsymbol{Z^Y} \times \boldsymbol{Y}$.

Here, we'll stop and look at a pretty cool real world example.

---

**Example 4.5.3.** Consider the category **Set**. Then we know that, for any two given objects $Y$ and $Z$, we can form a set of functions between the objects:

$$\text{Hom}_{\textbf{Set}}(Y, Z).$$

Thus, the collection of morphisms from sets $Y$ to $Z$ is *itself a set*, and hence a member of **Set**. Now let $A$ be any object in **Set**, and let

$$X = \{f \in \textbf{Set} \mid f : A \times Y \longrightarrow Z\}.$$

Define **eval** : $\text{Hom}_{\textbf{Set}}(Y, Z) \times Y \longrightarrow Z$ as, who would've guessed, the evaluation:

$$\textbf{eval}(f(y), y') = f(y').$$

Now for each $a \in A$, we can define a function $g_a : X \times Y \longrightarrow Z$ where for each $f : A \times Y \longrightarrow Z$

$$g_a(f, y') = f(a, y') \in Z$$

so this is sort of a "double" evaluation function. Then for every such $g_a$, there exists a unique $h_a : X \longrightarrow \text{Hom}_{\textbf{Set}}(Y, Z)$ where for each $f : A \times Y \longrightarrow Z$

$$h_a(f) = f(a, y) : Y \longrightarrow Z.$$

Thus we get the following commutative diagram:

$$\begin{tikzcd}
Z & \text{Hom}_{\textbf{Set}}(Y,Z) \times Y \arrow[l, "\textbf{eval}"'] & & Z^Y \\
& X \times Y \arrow[ul, "g_a"] \arrow[u, dashed, "(h_a, \text{id}_Y)"'] & & X \arrow[u, dashed, "h_a"']
\end{tikzcd}$$

What is this? What's really going on and why do we care?
This construction relates to a concept in computer science called **currying**. Applied category theory in computer science generally works in **Set**, so that's why this idea transfers over.

The idea is: given a multivariable function, do we evaluate all arguments at once, or evaluate just one argument, thereby sending a function to another function? Both methods can offer advantages. But universality tells us that, in the end, they're the same thing.

We can think of $X \times Y$ as being elements $(f(a, y), y')$ where $f : A \times Y \longrightarrow Z$. Then $h$ evaluates $f(a', y)$ for some $a'$, thus sending the function $f : A \times Y \longrightarrow Z$ to the function $f : Y \longrightarrow Z$.

That is,

$$(h \times \mathrm{id}_y) \circ \Big((f(a,y), y')\Big) = (f(a',y), y').$$

Finally, **eval** evaluates $f(a', y)$ at $y'$, returning an object in $Z$.

Alternatively, we can start with the object $(f(a,y), y')$, and simply act on $g$, which evaluates it at both $a'$ and $y'$, returning the same object $f(a', y')$. Thus in the realm of computer science, we may think of the morphisms $(h, \mathrm{id}_y)$, $g$ and **eval** as commands, as this is how currying is often done.

The universality of this constructions states that both methods are the same; that is,

$$g = \textbf{eval} \circ (h \times \mathrm{id}_Y).$$

Since we started with arbitrary objects in **Set**, the consequence for computer science is that we can always curry these functions. Typically what is curried are types, such as **Bool** or **Int**.

---

In an arbitrary category of finite products, the exponential object is just a generalization of currying. But in **Set**, we see that an exponential object exists for any two pairs of sets. Thus, can we turn this exponential assignment into a functor? Yes,we can.

**Definition 4.5.4.** Let $\mathcal{C}$ have finite products and exponential objects for every pair of objects. Then for each $Y$ in $\mathcal{C}$ we can create an **exponential functor** $E^Y : \mathcal{C} \longrightarrow \mathcal{C}$ as follows.

**Objects.** For each $Z \in \mathcal{C}$, we define $E^Y(Z) = Z^Y$.

**Morphisms.** Let $f : A \longrightarrow B$ be in $\mathcal{C}$. Then we note that we have the following diagrams.

$$\begin{tikzcd}
A & A^Y \times Y \arrow[l, "\textbf{eval}_A"'] & A^Y \\
& X \times Y \arrow[ul, "g"] \arrow[u, dashed, "(h{,}\mathrm{id}_Y)"'] & X \arrow[u, dashed, "h"']
\end{tikzcd}
\qquad
\begin{tikzcd}
B & B^Y \times Y \arrow[l, "\textbf{eval}_B"'] & B^Y \\
& X \times Y \arrow[ul, "g'"] \arrow[u, dashed, "(h'{,}\mathrm{id}_Y)"'] & X \arrow[u, dashed, "h'"']
\end{tikzcd}$$

Now observe that we can form the morphism $f \circ \textbf{eval}_A : A^Y \times Y \longrightarrow B$. Hence by universality of $B^Y$, there exists a unique morphism $h' : A^Y \longrightarrow B^Y$. Diagrammatically, we take the above diagram on the right, and replace $X$ with $A^Y$ and $g$ with $f \circ \textbf{eval}_A$.

$$\begin{tikzcd}
B & B^Y \times Y \arrow[l, "\textbf{eval}_B"'] & B^Y \\
& A^Y \times Y \arrow[ul, "f \circ \textbf{eval}_A"] \arrow[u, dashed, "(h'{,}\mathrm{id}_Y)"'] & A^Y \arrow[u, dashed, "h'"']
\end{tikzcd}$$

Since $h$ exists if $f : A \longrightarrow B$ exists, we therefore define

$$E^Y(f : A \longrightarrow B) = h' : A^Y \longrightarrow B^Y$$

where $h'$ is the unique morphism such that

$$f \circ \mathbf{eval}_A = \mathbf{eval}_A \circ (h', \mathrm{id}_Y).$$

Note that there's one more cool connection here. If we have a category with finite products, and one in which exponential objects exist, then we have a morphism $\mathbf{eval}_A : A^Y \times Y \longrightarrow A$ which is universal from the functor $- \times Y : \mathcal{C} \longrightarrow \mathcal{C}$ to $A$. **Therefore, this is a counit!** There's an adjunction hiding here.

**Proposition 4.5.5.** Let $\mathcal{C}$ be a category with finite products and exponential objects. Let $Y$ be an object, and define the functors

$$P_Y = (-) \times Y : \mathcal{C} \longrightarrow \mathcal{C}$$
$$E^Y = (-)^Y : \mathcal{C} \longrightarrow \mathcal{C}.$$

Then $E^Y$ is right adjoint to $P_Y$ for every $Y \in \mathcal{C}$. Therefore,

$$\mathrm{Hom}_{\mathcal{C}}(X \times Y, Z) \cong \mathrm{Hom}_{\mathcal{C}}(X, Z^Y)$$

which is natural for all objects $X, Y, Z \in \mathcal{C}$.

*Proof.* For each object $A \in \mathcal{C}$, the exponential object gives rise to a universal morphism $\mathbf{eval}_A : A^Y \times Y \longrightarrow A$. So on one hand, we get the diagram on the left

$$\begin{tikzcd} A & A^Y \times Y \arrow[l, "\mathbf{eval}_A"'] & A^Y \\ & X \times Y \arrow[ul, "g"] \arrow[u, dashed, "(h, \mathrm{id}_Y)"'] & X \arrow[u, dashed, "h"'] \end{tikzcd} \qquad \begin{tikzcd} A & P_Y(E^Y(A)) \arrow[l, "\mathbf{eval}_A"'] & A^Y \\ & P_Y(X) \arrow[ul, "g"] \arrow[u, dashed, "(h, \mathrm{id}_Y)"'] & X \arrow[u, dashed, "h"'] \end{tikzcd}$$

but on the other hand, the diagram on the right is exactly equivalent. Hence we see that **eval** is actually a counit $\varepsilon_A : P_Y(E^Y(A)) \longrightarrow A$. Since such a counit exists for each $A$, this gives rise to an adjunction, so that $E^Y$ is right adjoint to $P_Y$ for every object $Y$ in $\mathcal{C}$. ∎

Finally, we have everything we need to move onto to the main point of this section.

**Definition 4.5.6.** Let $\mathcal{C}$ be a category. We say $\mathcal{C}$ is a **cartesian closed category** if the functors

$$U : \mathbf{C} \longrightarrow \mathbf{1} \qquad \Delta : \mathcal{C} \longrightarrow \mathcal{C} \times \mathcal{C} \qquad P_Y = (-) \times Y : \mathcal{C} \longrightarrow \mathcal{C}$$

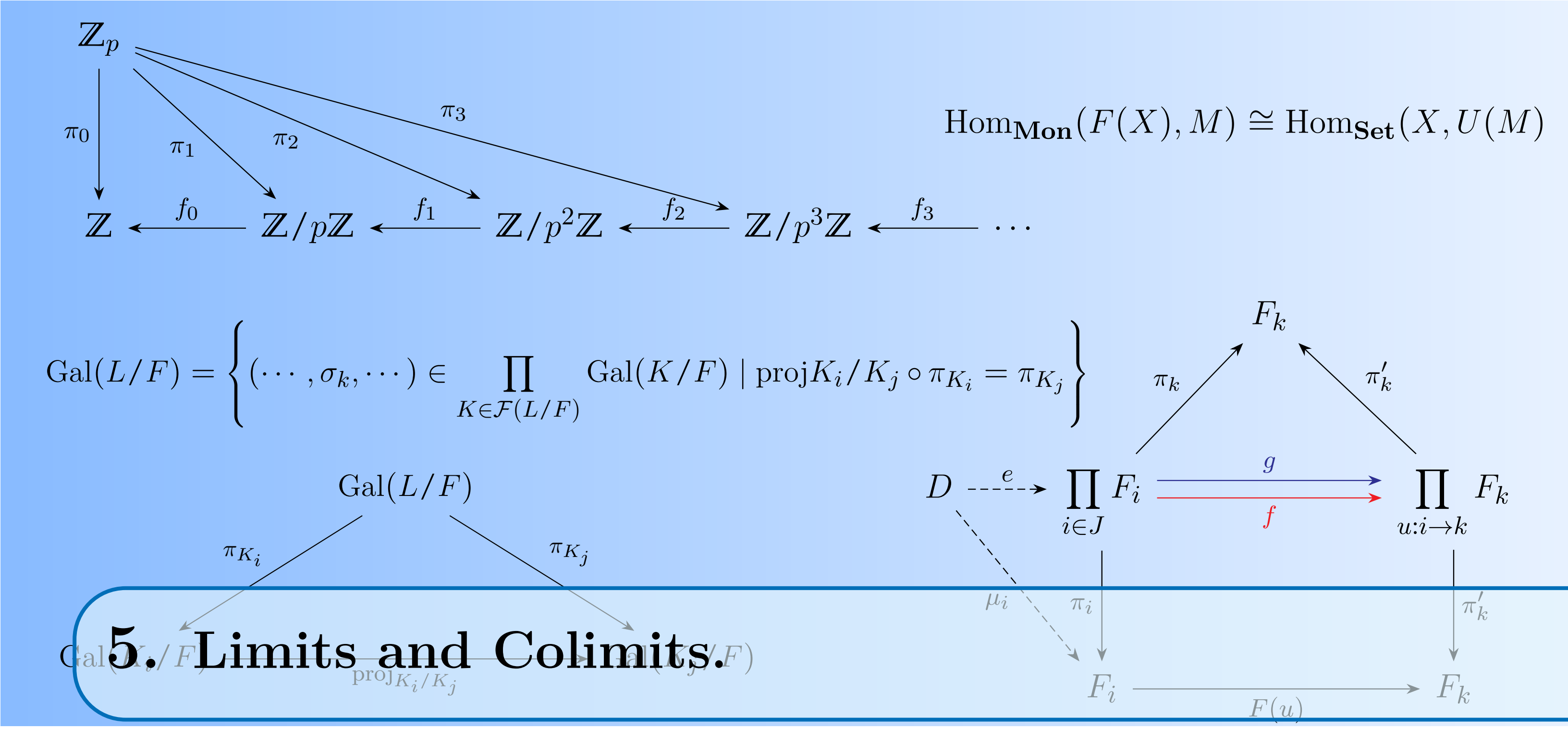
have right adjoints. In other words, $\mathcal{C}$ is **cartesian closed** if

1. There exists a terminal object $T$
2. $\mathcal{C}$ has finite products
3. An exponential object $A^Y$ for every $A \in \mathcal{C}$ for all $Y$.

Thus the work we just did was used in showing that our three-bullet point list is another definition of a cartesian closed category. Often, only one definition or the other is offered, and it's not trivial how they're equivalent, so it can be confusing. Thus our work shows that either definition is equivalent.

Some examples include **Set**, which we already dealt with. **Set** has a terminal object (empty set), has finite products, and has an exponential object. More interesting is **Cat**, which is cartesian closed. In this case, **1** is the terminal object, **Cat** is closed under finite products, and the exponential object exists. In this case, $\mathcal{C}^{\mathcal{B}}$ is simply the functor category!

At first, it seemed silly to define $\mathcal{C}^{\mathcal{B}}$ as the category of functors *from* $\mathcal{B}$ to $\mathcal{C}$, since it seemed that it ought to be denoted $\mathcal{B}^{\mathcal{C}}$. However, we see that this was really just because of the concept of exponentials, which isn't known when being introduced functor categories.




# 5. Limits and Colimits.

Before we begin, we reintroduce certain terminology.

**Definition 5.0.1.** Let $\mathcal{C}$ be a category. We define a **diagram** of a **shape** $J$ to be a functor $F : J \longrightarrow \mathcal{C}$.
Here, $J$ is generally thought of as an indexing category. We use the word diagram because the image of $J$ under $F$ is literally a diagram of morphisms.

$$
\begin{tikzcd}
 & i \arrow[dr, "g"] & \\
j \arrow[ur, "f"] \arrow[rr, "h"'] & & k
\end{tikzcd}
\quad \xrightarrow{F} \quad
\begin{tikzcd}
 & F(i) \arrow[dr, "F(g)"] & \\
F(j) \arrow[ur, "F(f)"] \arrow[rr, "F(h)"'] & & F(k)
\end{tikzcd}
$$

In this example, on the left we have the category $J$, and on the right we have the diagram of $J$ in $\mathcal{C}$. Now recall the **diagonal functor**

$$\Delta : \mathcal{C} \longrightarrow \mathcal{C}^J$$

is a functor which sends each object $C \in \mathcal{C}$ to the functor $\Delta(C) : J \longrightarrow \mathcal{C}$, where, for each $j \in J$, we have

$$\Delta(C)(j) = C.$$

This motivates the following concept.

**Definition 5.0.2.** Let $\mathcal{C}$ be a category and $F : J \longrightarrow \mathcal{C}$ be a functor, where $J$ is a small category. We define a **cone over $F$ with apex $C$** to be a natural transformation

$$\Delta(C) \longrightarrow F.$$

**Equivalently**, it is an object $C$ equipped with morphisms $u_i : C \longrightarrow F(i)$ for each $i \in J$ such that, for every $f : i \longrightarrow j$ in $J$, the diagram

$$\begin{tikzcd}
& C \arrow[dl, "u_i"'] \arrow[dr, "u_j"] & \\
F(i) \arrow[rr, "f"'] & & F(j)
\end{tikzcd}$$

commutes.

In the same fashion, we may define a **cocone with base $C$ under $F$** as a natural transformation

$$F \longrightarrow \Delta(C).$$

**Equivalently**, it is an object $C$ equipped with morphisms $u_i : F(i) \longrightarrow C$ for each $i \in J$ such that, for every $f : i \longrightarrow j$ in $J$, the diagram

$$\begin{tikzcd}
F(i) \arrow[rr, "f"] \arrow[dr, "u_i"'] & & F(j) \arrow[dl, "u_j"] \\
& C &
\end{tikzcd}$$

commutes.

Alternatively, we could have defined the above, second definition as a "cone," and then defined the first definition as the "cocone". Why? Well, it's the same arbitrary nature in which physicists encountered electrical charge; one was named negative, the other was named positive. For all we know, in a parallel universe protons were said to have "negative" charge and electrons were said to have "positive" charge. In the end, nomenclature is arbitrary when it comes to duality.

Try to recall: what is a **limit** in a category $\mathcal{C}$? When we speak of one, we're talking about the limit of a functor

$$F : J \longrightarrow \mathcal{C}.$$

There are multiple, but equivalent ways to think about it.

- A limit can be thought of as a **universal object** $(\operatorname{Lim} F, u : \Delta(\operatorname{Lim} F) \longrightarrow F)$ from $\Delta$ to $F$.

$$\begin{tikzcd}
F & \Delta(\operatorname{Lim} F) \arrow[l, "u"'] & & \operatorname{Lim} F \\
& \Delta(C) \arrow[ul, "f"] \arrow[u, dashed, "\Delta(h)"'] & & C \arrow[u, dashed, "h"']
\end{tikzcd}$$

- A limit can also be thought of as a **universal cone**. We know that if we have a limit, then we have an object $\mathrm{Lim}\, F$ and a natural transformation $\Delta(\mathrm{Lim}\, F) \longrightarrow F$. Hence, this forms a cone. As we also pointed out, a cone induces a family of morphisms $u_i : \mathrm{Lim}\, F \longrightarrow F(i)$.

  What makes this cone a "universal" cone is the fact that, for any other cone $\Delta(C) \longrightarrow F$, the above diagram establishes the diagram below.

$$\begin{tikzcd}
& C \arrow[d, dashed, "h"] \arrow[ddl, bend right, "f_i"'] \arrow[ddr, bend left, "f_j"] & \\
& \mathrm{Lim}\, F \arrow[dl, "u_i"'] \arrow[dr, "u_j"] & \\
F(i) \arrow[rr, "F(g)"] & & F(j)
\end{tikzcd}$$

- In a better way, one can think of it as being a *universal spider!* One could also think of it as a squished spider, or more optimistically, a two dimensional spider.

$$\begin{tikzcd}
& & C \arrow[d, dashed, "f"] \arrow[ddll, bend right, "v_i"'] \arrow[ddl, bend right, "v_j"'] \arrow[ddr, bend left, "v_k"] \arrow[ddrr, bend left, "v_l"] & & \\
& & \mathrm{Lim}\, F \arrow[dll, "u_i"'] \arrow[dl, "u_j"] \arrow[dr, "u_k"'] \arrow[drr, "u_l"] & & \\
F(i) & F(j) & \cdots & F(k) & F(l)
\end{tikzcd}$$

Now try to recall what **Colimits** of a diagram $F : J \longrightarrow \mathcal{C}$ are. As before, there are multiple, but equivalent ways to think about it.

- A colimit can be thought of as a **universal object** $(\mathrm{Colim}\, F, u : F \longrightarrow \Delta(\mathrm{Colim}\, F))$ from $F$ to $\Delta$.

$$\begin{tikzcd}
F \arrow[r, "u"] \arrow[dr, "f"'] & \Delta(\mathrm{Colim}\, F) \arrow[d, dashed, "\Delta(h)"] & \mathrm{Lim}\, F \arrow[d, dashed, "h"'] \\
& \Delta(C) & C
\end{tikzcd}$$

- A colimit can also be thought of as a **universal cocone** (or just cone). Given a colimit of $F : J \longrightarrow \mathcal{C}$, we have an object $\mathrm{Colim}\, F$ and a natural transformation $u : F \longrightarrow \Delta(\mathrm{Colim}\, F)$. Hence, this forms a cocone (again, or just cone). As we also pointed out, a cocone induces a family of morphisms $u_i : F(i) \longrightarrow \mathrm{Colim}\, F$.

What makes this cocone a "universal" cocone is the fact that, for any other cocone $F \longrightarrow \Delta(C)$, the above diagram establishes the diagram below.

$$\begin{tikzcd}
F(i) \arrow[rr, "F(g)"] \arrow[dr, "u_i"'] \arrow[ddr, bend right, "f_i"'] & & F(j) \arrow[dl, "u_j"] \arrow[ddl, bend left, "f_j"] \\
& \operatorname{Lim} F \arrow[d, dashed, "h"] & \\
& C &
\end{tikzcd}$$

- One can also think of this as a *universal spider!* Or it can be thought of as a *jealous object*; if any other object $C$ is "the center of attention,"i.e. has morphisms pointing to it, $\operatorname{Colim} F$ will get angry, so the morphisms have to go through $\operatorname{Colim} F$ via $f$ first before they reach $C$.

$$\begin{tikzcd}
& & C & & \\
& & \operatorname{Colim} F \arrow[u, dashed, "h"'] & & \\
F(i) \arrow[urr, "u_i"] \arrow[uurr, bend left, "f_i"] & F(j) \arrow[ur, "u_j"'] \arrow[uur, bend left, "f_j"] & \cdots & F(k) \arrow[ul, "u_k"] \arrow[uul, bend right, "f_k"'] & F(l) \arrow[ull, "u_l"'] \arrow[uull, bend right, "f_l"']
\end{tikzcd}$$

# 5.1 Every Limit in Set; Creation of Limits

While we have been discussing limits and colimits of functors, we generally consider the case in which they exist. However, they sometimes don't exist; after all, limits and colimits are universal objects. Categories which do admit these constructions are often convenient places to work inside of. This is analogous to **complete metric spaces** $X$, where every Cauchy sequence is convergent in $X$. With such an analogy in mind, the following definition should make sense.

**Definition 5.1.1.** Let $\mathcal{C}$ be a category. We say $\mathcal{C}$ is **complete** if all small diagrams in $\mathcal{C}$ has limits in $\mathcal{C}$; in other words, if every functor $F : J \longrightarrow \mathcal{C}$, where $J$ is a small category, has a limit in $\mathcal{C}$.

Similarly, we define:

**Definition 5.1.2.** Let $\mathcal{C}$ be a category. We say $\mathcal{C}$ is **cocomplete** if all small diagrams in $\mathcal{C}$ has colimits in $\mathcal{C}$. In other words, every functor $F : J \longrightarrow \mathcal{C}$, where $J$ is a small category, has a colimit in $\mathcal{C}$.

Now we show how to construct limits inside of **Set**, thereby showing that this category is complete.

---

**Example 5.1.3.** For this example, let $J = \omega^{\text{op}}$, where $\omega$ is the preorder of natural numbers. Since we are asking for the opposite category, we reverse the arrows and get the diagram below.

$$\begin{tikzcd}
0 \arrow[loop above] & 1 \arrow[l] \arrow[loop above] & 2 \arrow[l] \arrow[loop above] & 3 \arrow[l] \arrow[loop above] & \cdots \arrow[l]
\end{tikzcd}$$

Now suppose $F : \omega^{\text{op}} \longrightarrow \textbf{Set}$ is a functor. Then if we write $F(i) = A_i$ with $A_i \in \textbf{Set}$, then we see that the image of $F$ is a family of sets $F_n$ with functions $f_n : A_{n+1} \longrightarrow A_n$:

$$\begin{tikzcd}
A_0 & A_1 \arrow[l, "f_0"'] & A_2 \arrow[l, "f_1"'] & A_3 \arrow[l, "f_2"'] & \cdots \arrow[l, "f_3"']
\end{tikzcd}$$

One way we could try forming a limit of this diagram is by constructing a cone, using the product of these sets.

$$\begin{tikzcd}
A_0 & A_1 \arrow[l, "f_0"'] & A_2 \arrow[l, "f_1"'] & A_3 \arrow[l, "f_2"'] & \cdots \arrow[l, "f_3"'] \\
\prod_{i=0}^{\infty} A_i \arrow[u, "\pi_0"] \arrow[ur, "\pi_1"] \arrow[urr, "\pi_2"] \arrow[urrr, "\pi_3"'] & & & &
\end{tikzcd}$$

However, this isn't exactly what we want. A cone must form a commutative diagram and it's not always true that $f_n \circ \pi_{n+1} = \pi_n$. So let's instead restrict our attention to a subset $L \subseteq \prod_{i=0} A_i$

where the points $(a_0, a_1, \dots, a_n, \dots)$ do satisfy this relation.

$$L = \Big\{x = (a_0, a_1, a_2, \dots,) \mid f_n \circ \pi_{n+1}(a_{}) = \pi_n(x)\Big\}.$$

and equip $L$ with the functions $\pi'_n$ where

$$\pi'_n = \pi_n \circ i : L \longrightarrow A_n$$

where $i : L \longrightarrow \prod_{i=1}^{\infty} F_i$ is the inclusion function. Then we have

$$\begin{tikzcd}
A_0 & A_1 \arrow[l, "f_0"'] & A_2 \arrow[l, "f_1"'] & A_3 \arrow[l, "f_2"'] & \cdots \arrow[l, "f_3"'] \\
L \arrow[u, "\pi'_0"] \arrow[ur, "\pi'_1"] \arrow[urr, "\pi'_2"] \arrow[urrr, "\pi'_3"'] & & & &
\end{tikzcd}$$

so that $L$ forms a cone. We now prove that this cone is universal.

**Lemma 5.1.4.** The set $L$ is the limit of the functor $F : \omega^{\text{op}} \longrightarrow \textbf{Set}$.

*Proof.* Suppose $K$ is another cone over our diagram, equipped with morphisms $\mu_n : K \longrightarrow F_n$. Since this is another cone, we have that $f_n \circ \mu_{n+1} = \mu_n$. Now let $k \in K$. Then we can form an element

$$x = (\mu_0(k), \mu_1(k), \mu_2(k), \dots) \in \prod_{i=1}^{\infty} F_i$$

since each $\mu_n(k) \in F_n$. Now observe that

$$f_n \circ \pi_{n+1}(x) = f_n(\mu_{n+1}(k)) = \mu_n(k) = \pi_n(x).$$

Thus we see that $f_n \circ \pi_{n+1}(x) = \pi_n(x)$, so that by definition, $x \in L$. Hence we can create a unique function $g : K \longrightarrow L$ where for each $k \in K$,

$$g(k) = (\mu_0(k), \mu_1(k), \mu_2(k), \dots)$$

so we then have that

$$\pi'_n \circ g = \mu_n.$$

Hence, this shows that $(L, \pi_n : L \longrightarrow F_n)$ is universal, so that $L = \text{Lim}\, F$!

$$\begin{tikzcd}
F_0 & F_1 \arrow[l, "f_0"'] & F_2 \arrow[l, "f_1"'] & F_3 \arrow[l, "f_2"'] & \cdots \arrow[l, "f_3"'] \\
L = \text{Lim}\, F \arrow[u] \arrow[ur] \arrow[urr] \arrow[urrr] & & & K \arrow[lll, dashed, "g"] \arrow[u] \arrow[ul] \arrow[ull] \arrow[ulll] &
\end{tikzcd}$$

■

If we want to view this in terms of the spider diagrams, then we have

$$\begin{tikzcd}
 & & K \arrow[ddll, bend right, "\mu_0"'] \arrow[ddl, bend right, "\mu_1"'] \arrow[d, dashed, "g"] \arrow[ddr, bend left, "\mu_2"] \arrow[ddrr, bend left, "\mu_3"] & & & \\
 & & L = \operatorname{Lim} F \arrow[dll, "\pi'_0"'] \arrow[dl, "\pi'_1"] \arrow[dr, "\pi'_2"'] \arrow[drr, "\pi'_3"] & & & \\
F_0 & F_1 \arrow[l, "f_0"] & & F_2 \arrow[ll, "f_1"] & F_3 \arrow[l, "f_2"] & \cdots \arrow[l, "f_3"]
\end{tikzcd}$$

---

Here, we've taken a nice, simple diagram $F : \omega^{\text{op}} \longrightarrow \textbf{Set}$ and shown that there exists a limit $L$ of the diagram inside of **Set**. However, we can do this more generally, so that **Set** is complete. To illustrate this we need the notion of a set of cones.

Note that in the last example, we can actually think of each $x = (x_0, x_1, x_2, \dots) \in \operatorname{Lim} F$ as a cone. How so?

1. For each $x = (x_0, x_1, x_2, \dots) \in \operatorname{Lim} F$, consider the one-point set $\{*\}$.
2. Associate $\{*\}$ with the family of functions $\pi_n^* : \{*\} \longrightarrow F_n$, defined as

$$\pi_n^x(*) = x_n.$$

Now since $x \in \operatorname{Lim} F$, we know that $f_n(x_{n+1}) = x_n$. But, note that this is equivalent to stating that $f_n \circ \pi_{n+1}(*) = \pi_{n-1}(*)$. Therefore the diagram

$$\begin{tikzcd}
 & \{*\} \arrow[dl, "\pi_n^x"'] \arrow[dr, "\pi_{n+1}^x"] & \\
F_n & & F_{n+1} \arrow[ll, "f_n"]
\end{tikzcd}$$

commutes for every $f_n : F_{n+1} \longrightarrow F_n$, so that's how we can regard every $x \in \operatorname{Lim} F$ as a cone. Therefore, if we denote $\text{Cone}(*, F)$ as the set of all cones of $\{*\}$ over $F$, we see that $\text{Cone}(*, F) = \operatorname{Lim} F$.

**Theorem 5.1.5.** The category **Set** is complete. That is, if $J$ is a small category, every functor $F : J \longrightarrow \mathbf{Set}$ has a limit

$$\text{Lim}\, F = \text{Cone}(*, F)$$

where $\text{Cone}(*, F)$ is the set of all cones of $\{*\}$ over $F$. The set $\text{Cone}(*, F)$ forms the limit cone with the morphisms $v_i : \text{Cone}(*, F) \longrightarrow F_i$ described as follows. If $x \in \text{Cone}(*, F)$, then $x$ has a family of morphisms $\sigma_i^x : \{*\} \longrightarrow F_i$. Therefore,

$$v_i : \text{Cone}(*, F) \longrightarrow F_i \qquad v_i(x) = \sigma_i^x(*).$$

*Proof.* First, since $J$ is small, we know that $\text{Cone}(*, F)$ is a set. For each $j \in J$, establish the morphism $v_j : \text{Cone}(*, F) \longrightarrow F_j$ where $v_j(x) = \sigma_j^x(x)$, and $\sigma_j^x : \{*\} \longrightarrow F_j$ is the morphism associated with $x$ as a cone over $F$.

We now show that it is a cone. Suppose $f : i \longrightarrow j$ is a morphism in $J$. Then observe that $F(f) \circ v_i(x) = F(f) \circ \sigma_i^x(x) = \sigma_j^x(x) = v_j(x)$. Hence the diagram

$$\begin{tikzcd}
 & \text{Cone}(*, F) \arrow[dl, "v_i"'] \arrow[dr, "v_j"] & \\
F_i \arrow[rr, "F(f)"'] & & F_j
\end{tikzcd}$$

commutes, so $\text{Cone}(*, F)$ really does form a cone over $F$. To show this is universal, and hence our limit, suppose that $A$ in **Set** also forms a cone over $F$ with morphisms $\tau_j : X \longrightarrow F_j$. Note that for each $a \in A$, we can form a cone from $\{*\}$ to $F$, if we define $\sigma_j^a : \{*\} \longrightarrow F_j$ as $\sigma_j^a(*) = \tau_j(a)$. Then the diagram

$$\begin{tikzcd}
 & \text{Cone}(*, F) \arrow[dl, "\sigma_i^a"'] \arrow[dr, "\sigma_j^a"] & \\
F_i \arrow[rr, "F(f)"'] & & F_j
\end{tikzcd}$$

must also commutes since it commutes for each $\tau_j$. Thus we can define a unique function $g : A \longrightarrow \text{Cone}(*, F)$, where each point $a$ is sent to the cone which it forms from $\{*\}$ over $F$. Therefore, $\text{Cone}(*, F)$ is universal, so that

$$\text{Lim}\, F = \text{Cone}(*, F)$$

as desired. ∎

The above proof can be repeated to show that others categories are complete, like **Grp** or **Rng**.

In attempting to find the limit $F : J \longrightarrow \mathcal{C}$ in some category $\mathcal{C}$, one strategy is to to compose this functor with another one $G : \mathcal{C} \longrightarrow \mathcal{D}$, with the prior knowledge that $\mathcal{D}$ is complete. If one knows $\mathcal{D}$ is complete, one then use this information to find the limit of $F : J \longrightarrow \mathcal{C}$.

**Definition 5.1.6.** Let $F : J \longrightarrow \mathcal{C}$ be a functor. A functor $G : \mathcal{C} \longrightarrow \mathcal{D}$ **creates limits for** $F$ if whenever $(\operatorname{Lim} G \circ F, \tau : \Delta(\operatorname{Lim} G \circ F) \longrightarrow G \circ F)$ exists, the limit $(\operatorname{Lim} F, \sigma : \Delta(\operatorname{Lim} F) \longrightarrow F)$ such that

$$G(\operatorname{Lim} F) = \operatorname{Lim} G \circ F \qquad G(\sigma) = \tau.$$

Similarly, a functor $G : \mathcal{C} \longrightarrow \mathcal{D}$ **creates colimits for** $F$ if whenever $(\operatorname{Colim} G \circ F, \tau : G \circ F \longrightarrow \Delta(\operatorname{Lim} G \circ F)$ exists, the colimit $(\operatorname{Colim} F, \sigma : F \longrightarrow \Delta(\operatorname{Colim} F)$ exists and

$$G(\operatorname{Colim} F) = \operatorname{Colim} G \circ F \qquad G(\sigma) = \tau.$$

The diagram below visually explains this process; the existence of limit in $\mathcal{D}$ on the left implies the existence of the limit in $\mathcal{C}$ on the right. Moreover, the diagram on the left is the image of the diagram on the right under $G$.

$$\begin{tikzcd}
& D \arrow[d, dashed, "h"] \arrow[ddl, bend right, "f_i"'] \arrow[ddr, bend left, "f_j"] & \\
& \operatorname{Lim} G \circ F \arrow[dl, "\tau_i"'] \arrow[dr, "\tau_j"] & \\
G(F(i)) \arrow[rr, "G(F(u))"'] & & G(F(j))
\end{tikzcd}
\quad \Longrightarrow \quad
\begin{tikzcd}
& C \arrow[d, dashed, "h"] \arrow[ddl, bend right, "f_i"'] \arrow[ddr, bend left, "f_j"] & \\
& \operatorname{Lim} F \arrow[dl, "\sigma_i"'] \arrow[dr, "\sigma_j"] & \\
F(i) \arrow[rr, "F(u)"'] & & F(j)
\end{tikzcd}$$

---

**Example 5.1.7.** Consider a functor $F : J \longrightarrow \mathbf{Grp}$. We'll show that the forgetful functor $U : \mathbf{Grp} \longrightarrow \mathbf{Set}$ creates limits for **Grp**.

By the previous theorem, we know that $U \circ F; J \longrightarrow \mathbf{Set}$ must have a limit $\operatorname{Cone}(*, U \circ F)$ with the family of morphisms $v_i : \operatorname{Cone}(*, U \circ F) \longrightarrow U \circ F_i$. Now denote the set $\operatorname{Cone}(*, U \circ F)$ as $L$. Then we can endow $L$ with a group structure.

- For any $\sigma, \tau \in L$, we define $\sigma \times \tau$ to be the cone where $(\sigma \times \tau)_i = \sigma_i \cdot \tau_i$, where $\cdot$ is the product in $F_i$.
- For $\sigma \in L$, we define the inverse to be the function $\sigma^{-1}$ where $(\sigma^{-1})_i = \sigma_i^{-1}$, with the inverse being taken in $F_i$.

All we're really doing here is taking advantage of the fact that each $\sigma, \tau$ is really just a family of functions $\sigma_i, \tau_i : \{*\} \longrightarrow F_i$. Thus we're taking advantage of the group structure in each $F_i$.

Thus $L = \operatorname{Cone}(*, U \circ F)$ is a group, which then makes the family of morphisms $v_i$ :

Cone$(*, U \circ F)$ into a family of group homomorphisms. To show this, simply observe that

$$v_i(\sigma \times \tau) = (\sigma \times \tau)_i = \sigma_i \cdot \tau_i = v_i(\sigma) \cdot v_i(\tau).$$

Now we claim that the cone Cone$(*, U \circ F)$ with the morphisms $v_i : \text{Cone}(*, U \circ F) \longrightarrow F_i$ is universal. To show this, let $G$ be a group and suppose $G$ forms a cone over $F$ with morphisms $\varphi_i : G \longrightarrow F_i$. Then $U(G)$ forms a cone over Cone$(*, U \circ F)$ in **Set** with morphisms $U(\varphi_i) : U(G) \longrightarrow U(F_i)$.

Since we know Cone$(*, U \circ F)$ is a universal cone in **Set**, there exists a $h : U(G) \longrightarrow L$ such that $U(\varphi_i) = U(v_i) \circ h_i$. However, note that $h$ can be thought of as a group homomorphism. For any $g, g' \in G$, we have

$$\begin{aligned} h_i(g \cdot g') = \varphi_i(g \cdot g') = \varphi_i(g) \times \varphi_i(g') &= h_i(g) \times h_i(g') \\ &= (h(g) \cdot h(g'))_i. \end{aligned}$$

Therefore, $h : U(G) \longrightarrow L$ can be realized back into **Grp** as a group homomorphism $h : G \longrightarrow L$, thereby showing Cone$(*, U \circ F)$ is a universal cone in **Grp**. This is one way in showing **Grp** is complete.

---

What we really did in the last example was nothing special. Using the fact that **Set** is complete, we transferred $F : J \longrightarrow$ **Grp** over to **Set** via the forgetful functor $U :$ **Grp** $\longrightarrow$ **Set**. We calculated the limit, and showed that this can be realized as a limit in **Grp**. In this sense, $U :$ **Grp** $\longrightarrow$ **Set creates limits** in **Grp**. A similar strategy can be carried out for other forgetful functors.

---

**Example 5.1.8.** Let $\mathcal{C}$ be a category and $A$ an object of $\mathcal{C}$. Recall that with the comma category $(A \downarrow \mathcal{C})$, we have a projection functor $P : (A \downarrow \mathcal{C}) \longrightarrow \mathcal{C}$ where on objects $(C, f : A \longrightarrow C)$ and morphisms $h : (C, f : A \longrightarrow C) \longrightarrow (C', f : A \longrightarrow C')$ we have that

$$P(C, f : A \longrightarrow C) = C \qquad P(h) = h : C \longrightarrow C'.$$

Now for any functor $F : J \longrightarrow (A \downarrow \mathcal{C})$, the functor $P : (A \downarrow \mathcal{C}) \longrightarrow \mathcal{C}$ creates limits. To see this, we first interpret a functor $F : J \longrightarrow (A \downarrow \mathcal{C})$. For each $j$, we have that

$$F(j) = (C_j, f_j : A \longrightarrow C_j)$$

for some $C_j \in \mathcal{C}$ and $f_j : A \longrightarrow C_j$. If $u : j \longrightarrow k$ is a morphism in $J$, then $F(u) : C_j \longrightarrow C_k$ is a morphism in $\mathcal{C}$ such that the diagram below commutes (as, that's what morphisms do in comma categories).

$$\begin{tikzcd}
 & A \arrow[dl, "f_j"'] \arrow[dr, "f_k"] & \\
C_j \arrow[rr, "F(u)"'] & & C_k
\end{tikzcd}$$

Note that this is a cone over $F$ in $\mathcal{C}$. Now suppose we have a limit $\operatorname{Lim} P \circ F$ in $\mathcal{C}$ with morphisms $\mu_i : \operatorname{Lim} P \circ F \longrightarrow C_i$ with $i \in J$. Then because $\operatorname{Lim} P \circ F$ is a limiting cone, and we must have a unique $v$ such that the diagram below commutes.

$$\begin{tikzcd}
 & A \arrow[d, dashed, "v"] \arrow[ddl, bend right, "f_i"'] \arrow[ddr, bend left, "f_j"] & \\
 & \operatorname{Lim} P \circ F \arrow[dl, "\mu_i"'] \arrow[dr, "\mu_j"] & \\
F(i) \arrow[rr, "F(u)"'] & & F(j)
\end{tikzcd}$$

The claim now is that $(\operatorname{Lim} P \circ F, v : A \longrightarrow \operatorname{Lim} P \circ F)$ is the limit $\operatorname{Lim} F$ of $F : J \longrightarrow (A \downarrow \mathcal{C})$, which is left for the reader to show.

---

## Exercises

**1.** *i.* Let $J = \omega$, and let $F : J \longrightarrow \textbf{Set}$ be a functor were $F(i) = A_i$. Show that $\operatorname{Colim} F$ exists and give an expicit description of it.
*Hint*: It will be a set endowed with an equivalence relation.

*ii.* How does your answer chance when $F : J \longrightarrow \textbf{Set}$ is contravariant?

## 5.2 Inverse and Direct Limits.

In the previous example, we calculated the limit of the diagram indexed by $\omega^{\text{op}}$. It turns out that in general, we can construct a lot of mathematical ideas by first modeling them as the limit of a functor $F : J \longrightarrow \mathcal{C}$, where $J$ is a partially ordered set. Thus we give a special name to this concept.

**Definition 5.2.1.** Let $\mathcal{C}$ be a category, and suppose the $F : J^{\text{op}} \longrightarrow \mathcal{C}$ has a limit object $\text{Lim}\, F$ in $\mathcal{C}$, where $J$ is a partially ordered set (where, if $i \le j$, then there exists $f : i \longrightarrow j$). Then $\text{Lim}\, F$ is said to be a **inverse limit** or **projective limit**.

Dually, we define the colimit of a functor $F : J \longrightarrow F$ to be **direct limit**.

There are many famous examples of these limits, with the following example probably being the most familiar.

---

**Example 5.2.2.** Consider the functor $F : \omega^{\text{op}} \longrightarrow \textbf{Rng}$ where we define $F(n) = F_n = \mathbb{Z}/p^n\mathbb{Z}$ with $p$ being a prime. Then we have a diagram

$$\mathbb{Z} \xleftarrow{f_0} \mathbb{Z}/p\mathbb{Z} \xleftarrow{f_1} \mathbb{Z}/p^2\mathbb{Z} \xleftarrow{f_2} \mathbb{Z}/p^3\mathbb{Z} \xleftarrow{f_3} \cdots$$

where the maps $f_n : \mathbb{Z}/p^{n+1}\mathbb{Z} \longrightarrow \mathbb{Z}/p^n\mathbb{Z}$ are the projection maps. The limit of this diagram turns out to be the ***p*-adic integers** $\mathbb{Z}_p$, and this is one way of defining them. The most popular way to define them it to work in ring theory, establish $p$-adic valuations, and realize that the valuations turn $\mathbb{Z}$ into a metric space; one which can be completed with respect to the metric to give rise to $\mathbb{Z}_p$.

First, observe that they form a cone. Define the map

$$\pi_n : \mathbb{Z}_p \longrightarrow \mathbb{Z}/p^n\mathbb{Z} \qquad \pi\left(\sum_{k=0}^{\infty} a_k p^k\right) = \sum_{k=0}^{n-1} a_k p^k + p^n\mathbb{Z}.$$

Now observe that

$$\begin{aligned} f_n \circ \pi_{n+1}\left(\sum_{k=0}^{\infty} a_k p^k\right) = f_n\left(\sum_{k=0}^{n} a_k p^k + p^{n+1}\mathbb{Z}\right) &= \sum_{k=0}^{n-1} a_k p^k + p^n\mathbb{Z} \\ &= \pi_n\left(\sum_{k=0}^{\infty} a_k p^k\right) \end{aligned}$$

so we may conclude that $f_n \circ \pi n + 1 = \pi_n$. Therefore, $\mathbb{Z}_p$ does in fact form a cone with the morphisms $\pi_n$, so the following diagram commutes.

$$\begin{tikzcd} \mathbb{Z}_p \arrow[d, "\pi_0"'] \arrow[dr, "\pi_1"] \arrow[drr, "\pi_2"] \arrow[drrr, "\pi_3"] & & & & \\ \mathbb{Z} & \mathbb{Z}/p\mathbb{Z} \arrow[l, "f_0"'] & \mathbb{Z}/p^2\mathbb{Z} \arrow[l, "f_1"'] & \mathbb{Z}/p^3\mathbb{Z} \arrow[l, "f_2"'] & \cdots \arrow[l, "f_3"'] \end{tikzcd}$$

Showing this is universal is simple once we realize that each element of $\mathbb{Z}_p$ may be thought of as a cone, in the same fashion as we did with **Set**. That is, we can just apply the previous theorem to **Rng**. This then shows that it's the universal object which we desire.

---

What about direct limits? A less-talked about idea , although definitely not less interesting, is the dual of the above construction.

---

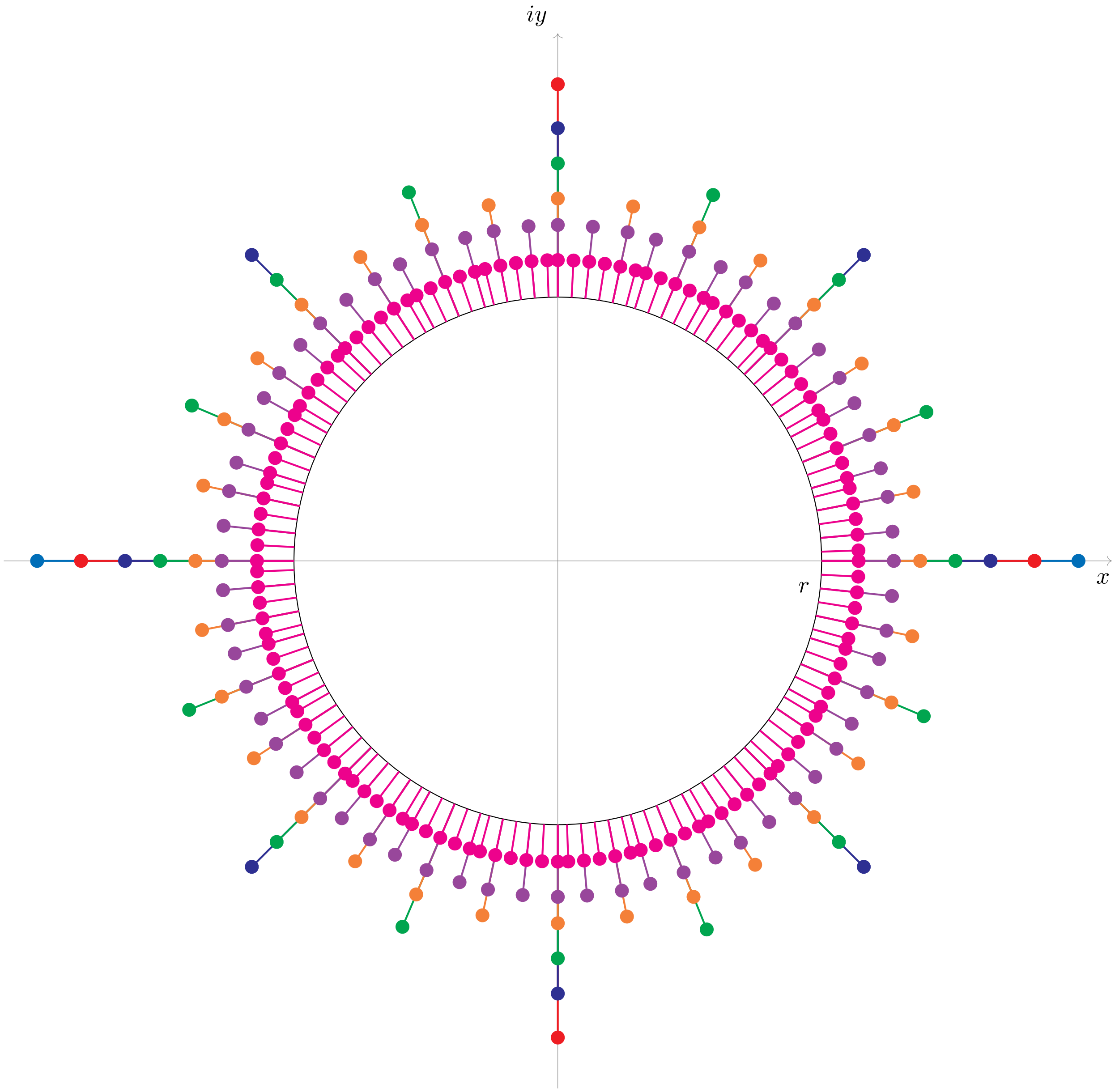
**Example 5.2.3.** Consider the functor $F : \omega \longrightarrow \textbf{Grp}$ where we have $F(n) = F_n = \mathbb{Z}/p^n\mathbb{Z}$, with $p$ being a prime. This time however we have the diagram

$$\mathbb{Z} \xrightarrow{f_0} \mathbb{Z}/p\mathbb{Z} \xrightarrow{f_1} \mathbb{Z}/p^2\mathbb{Z} \xrightarrow{f_2} \mathbb{Z}/p^3\mathbb{Z} \xrightarrow{f_3} \cdots$$

where we define each $f_n : \mathbb{Z}/p^n\mathbb{Z} \longrightarrow \mathbb{Z}/p^{n+1}\mathbb{Z}$ as the homomorphism

$$f_n\left(\sum_{k=0}^{n-1} a_k p^k + p^n\mathbb{Z}\right) = \sum_{k=0}^{n} a_k p^{k+1} + p^{n+1}\mathbb{Z}.$$

That is, we simply multiply the sum by a power of $p$. It turn outs that the direct limit is the **Prüfer $p$-Group** $\mathbb{Z}(p^\infty)$. The Prüfer 2-Group is pictured below.




The Prüfer $p$-group is the set of all $p^n$ roots of unity, as $n$ ranges over all positive integers. Hence the points lie on the complex unit circle. Specifically, it is the group

$$\mathbb{Z}(p^\infty) = \left\{ \exp\left(\frac{2\pi i m}{p^n}\right) \mid 0 \leq m < p^n, n \in \mathbb{Z}^+ \right\}$$

which forms a group under complex multiplication. How does this form a limit for our diagram?

---

Inverse limits are also used in Galois Theory. In Galois Theory, one can define a field extension $L/F$ to be a finite, normal, separable extension. However, it turns out that one can remove the requirement for the extension to be finite. We then obtain infinite Galois groups, which are constructed as follows.

---

**Example 5.2.4.** Let $F$ be a field, and suppose $L/F$ is normal, separable extension (not necessarily finite!). Then we can define $L/F$ to be a Galois extension, and we may speak of a Galois group $\text{Gal}(L/F)$, as follows.

Let $\mathcal{F}(L/F)$ be the category of all finite, normal extensions $K$ of $F$ such that $F \subseteq K \subseteq L$, and $\mathcal{G}(L/F)$ is the category of all their Galois groups. Note that both $\mathcal{F}(L/F)$ and $\mathcal{G}(L/F)$ are partially ordered sets, ordered by subset inclusion. To be precise, if $K_i \subseteq K_j$ are in $\mathcal{F}(L/F)$, then

$$\text{Gal}(K_j/F) \subseteq \text{Gal}(K_i/F)$$

and because $\mathcal{G}(L/F)$ is a preorder on subset inclusion, this implies the existence of some arrow $f : \text{Gal}(K_j/F) \longrightarrow \text{Gal}(K_i/F)$. We can describe $f = \text{proj}_{K_j/K_i}$ where

$$\text{proj}_{K_j/K_i} : \text{Gal}(K_j/F) \longrightarrow \text{Gal}(K_i/F) \qquad \text{proj}_{K_j/K_i}(\sigma) = \sigma\Big|_{K_i}.$$

That is, we take each permutation $\sigma \in \text{Gal}(K_j/F)$ and restrict its action to $K_i$, thereby making it a permutation of $K_i$ which fixes $F$, and therefore a member of $\text{Gal}(K_i/F)$.

Now consider the product with the associated morphisms

$$\prod_{K\in\mathcal{F}(L/F)} \text{Gal}(K/F) \qquad \pi_{K_i} : \prod_{K\in\mathcal{F}(L/F)} \text{Gal}(K/F) \longrightarrow \text{Gal}(K_i/F)$$

Then we define

$$\text{Gal}(L/F) = \left\{ x = (\cdots, \sigma_k, \cdots) \in \prod_{K\in\mathcal{F}(L/F)} \text{Gal}(K/F) \mid \text{proj}_{K_i/K_j} \circ \pi_{K_i}(x) = \pi_{K_j}(x) \right\}.$$

So $\text{Gal}(L/F)$ forms a cone with morphisms $\pi_{K_i}$:

$$\begin{tikzcd}
 & \text{Gal}(L/F) \arrow[dl, "\pi_{K_i}"'] \arrow[dr, "\pi_{K_j}"] & \\
\text{Gal}(K_i/F) \arrow[rr, "\text{proj}_{K_i/K_j}"'] & & \text{Gal}(K_j/F)
\end{tikzcd}$$

We then have to work to show that this cone is universal. However, the faster route is to simply recognize that we can index $\mathcal{G}(L/F)$ in a monotonic way, since it is a partially order set. Thus there exists a partially ordered set $J$ such that if $f : i \longrightarrow j$ exists in $J$, then

$$F(i) = \text{Gal}(K_i/F) \quad F(j) = \text{Gal}(K_j/F) \implies F(f) : \text{Gal}(K_i/F) \longrightarrow \text{Gal}(K_j/F)$$

Thus we have a functor $F : J \longrightarrow \mathcal{G}(L/F)$ which hits every Galois group $\text{Gal}(K/F)$ in such a way that it preserves the order in $\mathcal{G}(L/F)$. Since the limit of every small diagram exists in **Grp**, we can define $\text{Gal}(L/F)$ to be the **inverse limit** of this functor, and we already know that the limit will have the form

$$\text{Gal}(L/F) = \left\{ (\cdots, \sigma_k, \cdots) \in \prod_{K \in \mathcal{F}(L/F)} \text{Gal}(K/F) \mid \text{proj}_{K_i/K_j} \circ \pi_{K_i} = \pi_{K_j} \right\}.$$

and that it will be universal. So, this is how we extend the definition of Galois group from a finite, normal, separable extension to simple a normal, separable extension.

---

This construction can be done more generally on a partially ordered system of groups, to create these things called **profinite groups**.

**Definition 5.2.5.** Suppose we are given a partially ordered set of finite groups $G_i$, indexed by some set $I$, equipped with morphisms $\{f_i^j : G_j \longrightarrow G_i \mid i, j \in I \quad i \le j\}$ such that

1. $f_i^i : G_i \longrightarrow G_i$ is the identity $\text{id}_{G_i}$
2. $f_i^j \circ f_j^k = f_i^k$.

Then we define the **profinite group** $G$ of this system to be the inverse limit:

$$G = \left\{ (g_i)_{i \in I} \in \prod_{i \in I} G_i \mid f_i^j(g_i) = g_j \right\}.$$

Note that requiring $f_i^j(g_i) = g_j$ is the same as requiring $f_i^j \circ \pi_i(x) = \pi_j(x)$, where $x \in G$, which is how we defined $\text{Gal}(L/F)$.

Thus in the previous example, we have that not only can we actually define $\text{Gal}(L/F)$, but our construction leads to it to becoming a profinite group. Profinite groups are actually very special, in that they can be interpreted topologically.

## 5.3 Limits from Products, Equalizers, and Pullbacks.

In our construction of limits for **Sets**, we basically forced the existence of a cone, because we could. This is usually the general strategy when it comes to calculating the limit of a diagram in a given category; one uses available, useful constructions which are already present inside of a category. For example; in **Set**, we used the fact that it is cartesian closed to formulate infinite products.

Since the general strategy for showing **Set** is complete can be extended to other categories, one may wonder "well, why? And when will I no longer be able to apply this strategy?" The theorem below answers this question.

**Theorem 5.3.1.** Let $\mathcal{C}$ be a category and $J$ a small category. Suppose $\mathcal{C}$ has equalizers for every pair of morphisms in $\mathcal{C}$, and all products indexed by objects of $J$ and morphisms of $J$. Then every functor $F : J \longrightarrow \mathcal{C}$ has a limit in $\mathcal{C}$.

What do we mean by all products "indexed by objects of $J$ and morphisms of $J$"? What we want to do is be able to *create* products of the form

$$\prod_{j \in J} F_j \qquad\qquad \prod_{u:i \longrightarrow k} F_{\text{cod(u)}} = \prod_{u:j \longrightarrow k} F_k.$$

and *know* that they're in $\mathcal{C}$. The product on the far left is indexed by objects of $J$, while the equal ones on the right are indexed by morphisms $u : i \longrightarrow k$ in $J$. It's a bit weird to think of a product "indexed by morphisms," but it's exactly what it sounds like: we index over all the morphisms, and take the product of the domain or codomain (in the above, we did codomain).

Why do we need this weird concept? To answer this, let's go over the construction of limits in **Set** in a bit different way.

When we had a diagram $F : J \longrightarrow \mathcal{C}$ in $\mathcal{C}$, our first guess in constructing the limit was designing the $\prod_j F_j$ with morphisms $\pi_i : \prod_j F_j \longrightarrow F_i$. However, this doesn't actually form a cone, since for each $u : j \longrightarrow k$, we can't guarantee

$$F(u) \circ \pi_j = \pi_k$$

That is, we can't guarantee the diagram

$$\begin{tikzcd} & \prod_{j \in J} F_j \arrow[dl, "\pi_j"'] \arrow[dr, "\pi_k"] & \\ F_k \arrow[rr, "u"'] & & F_k \end{tikzcd}$$

will commute, which is what we need for a cone. Since we needed $F(u) \circ \pi_j = \pi_k$, we forced it. But this forcing is simply realizing that, all $x \in \prod_{j \in J} F_j$ which satisfy $F(u) \circ \pi_j = \pi_k$, are simply members of the equalizer of $F(u) \circ \pi_j$ and $\pi_k$.

*Proof.* Consider the products $\prod_{j \in J} F_j$ and $\prod_{u:i \longrightarrow k} F_k$ where in the last product we index over all morphisms in $J$. With both products, consider the projection morphisms

$$\pi'_j : \prod_{u:i \longrightarrow k} F_k \longrightarrow F_j$$
$$\pi_j : \prod_{i \in J} F_i \longrightarrow F_j.$$

Note that because we have products, we have universal properties which we can take advantage of. That is, the following diagrams must commute for some $f$ and $g$.

$$\begin{tikzcd}
 & \prod_{i \in J} F_i \arrow[dl, "\pi_k"'] \arrow[d, dashed, "f"] \\
F_k & \prod_{u:j \longrightarrow k} F_k \arrow[l, "\pi'_k"']
\end{tikzcd}
\qquad
\begin{tikzcd}
\prod_{i \in J} F_i \arrow[r, dashed, "g"] \arrow[d, "\pi_i"'] & \prod_{u:i \longrightarrow k} F_k \arrow[d, "\pi'_k"] \\
F_i \arrow[r, "F(u)"'] & F_k
\end{tikzcd}$$

Note however that we can stack these diagrams on top of each other, to obtain

$$\begin{tikzcd}
 & F_k & \\
\prod_{i \in J} F_i \arrow[ur, "\pi_k"] \arrow[rr, shift left, "g"] \arrow[rr, shift right, "f"'] \arrow[d, "\pi_i"'] & & \prod_{u:i \longrightarrow k} F_k \arrow[ul, "\pi'_k"'] \arrow[d, "\pi'_k"] \\
F_i \arrow[rr, "F(u)"'] & & F_k
\end{tikzcd}$$

Since we have equalizers for every pair of arrows, we can form the equalizer $e : D \longrightarrow \prod_{i \in J} F_i$ of both $f$ and $g$ for some object $D$.

$$\begin{tikzcd}
D \arrow[r, dashed, "e"] & \prod_{i \in J} F_i \arrow[r, shift left, "g"] \arrow[r, shift right, "f"'] & \prod_{u:i \longrightarrow k} F_k
\end{tikzcd}$$

Now that we have a morphism $e : D \longrightarrow \prod_{i \in J} F_i$, we can compose this with projections $\prod_{i \in J} F_i \longrightarrow F_i$ to produce a family of morphisms $\pi_i \circ e : D \longrightarrow F_i$. If we like, we can even add this to our diagram above to get the following:

$$\begin{tikzcd}
 & & F_k & \\
D \arrow[r, dashed, "e"] \arrow[dr, dashed, "\mu_i"'] & \prod\limits_{i \in J} F_i \arrow[ur, "\pi_k"] \arrow[rr, shift left, "g"] \arrow[rr, shift right, "f"'] \arrow[d, "\pi_i"] & & \prod\limits_{u:i \longrightarrow k} F_k \arrow[ul, "\pi'_k"'] \arrow[d, "\pi'_k"] \\
 & F_i \arrow[rr, "F(u)"'] & & F_k
\end{tikzcd}$$

(It looks like a boat!) Denote $\mu_i = \pi_i \circ e : D \longrightarrow F_i$. Then what the above boat diagram tells us is that

$$\pi'_k \circ g = \pi_k \qquad F(u) \circ \pi_i = \pi'_k \circ f.$$

Composing both equations with $e$, we get

$$\pi'_k \circ g \circ e = \pi_k \circ e \qquad F(u) \circ \pi_i \circ e = \pi'_k \circ f \circ e.$$

but since $g \circ e = f \circ e$, what this really tells us is that

$$F(u) \circ \pi_i \circ e = \pi_k \circ e \implies F(u) \circ \mu_i = \mu_k.$$

for every $u : i \longrightarrow k$ in $J$. Therefore, we see that we have that

$$\begin{tikzcd}
 & D \arrow[dl, "\mu_i"'] \arrow[dr, "\mu_j"] & \\
F_i \arrow[rr, "F(u)"'] & & F_k
\end{tikzcd}$$

commutes, so that $D$ equipped with the morphisms $\mu_i : D \longrightarrow F_i$ forms a cone. We now show that this is universal, so that $D$ is our limit. We do this by taking advantage of the universal property which equalizers posses.

Suppose $C$ is another object which forms a cone with morphisms $\tau_i : C \longrightarrow F_i$. Then there exists a map $e' : C \longrightarrow \prod\limits_{i \in J} F_i$ such that $\pi \circ e' = \tau_i$. Moreover, this implies that $f \circ e = g \circ e$. But the universal property of the equalizer $e$ states that for any subject object, there exists a morphism $h : D \longrightarrow C$ such that the diagram below commutes.

$$\begin{tikzcd}
D \arrow[r, "e"] \arrow[d, dashed, "h"'] & \prod\limits_{i \in J} F_i \arrow[r, shift left, "g"] \arrow[r, shift right, "f"'] \arrow[d, "\pi_i"] & \prod\limits_{u:i \longrightarrow k} F_k \\
C \arrow[ur, "e'"] \arrow[r, "\tau_i"] & F_i &
\end{tikzcd}$$

Since $h : D \longrightarrow C$ is unique, this shows that $D$ equipped with the morphisms $\mu_i : D \longrightarrow F_i$ forms a limit of the diagram, so that $D = \text{Lim}\, F$. ∎

We actually proved much more than what was stated in the theorem, since we literally found the explicit form the limit.

As a corollary, we have the following result which is due to the above theorem. The only difference is we strengthen our hypothesis, which makes it less general.

**Corollary 5.3.2.** Let $\mathcal{C}$ be a category. If $\mathcal{C}$ has all equalizers (coequalizers) and finite products (coproducts), then $\mathcal{C}$ has all finite limits (colimits).

By Proposition 3.3.8, one can obtain finite products by simply demanding the existence of binary products and a terminal object. Hence we can restate the above corollary:

**Corollary 5.3.3.** Let $\mathcal{C}$ be a category. If $\mathcal{C}$ has all equalizers (coequalizers), binary products (coproducts) and a terminal object, then $\mathcal{C}$ has all finite limits.

Not what is even more interesting is that we can construct equalizers and finite products from pullbacks.

Specifically, suppose our category $\mathcal{C}$ has pullbacks and a terminal object $T$. For any pair of objects $A, B$ in $\mathcal{C}$, suppose we take the pull back on the morphisms $t_A : A \longrightarrow T$ and $t_B : B \longrightarrow T$. This then give rise to an object $P$ equipped with two morphisms $p_1 : P \longrightarrow A$ and $p_2 : P \longrightarrow B$, universal in the sense demonstrated below.

$$
\begin{tikzcd}
C \arrow[dr, dashed, "h"] \arrow[drr, bend left, "f"] \arrow[ddr, bend right, "g"'] & & \\
& P \arrow[r, "p_1"] \arrow[d, "p_2"'] & A \arrow[d, "t_A"] \\
& B \arrow[r, "t_B"'] & T
\end{tikzcd}
\quad \Longrightarrow \quad
\begin{tikzcd}
& C \arrow[dl, "g"'] \arrow[dr, "f"] \arrow[d, dashed, "h"] & \\
B & P \arrow[l, "p_2"] \arrow[r, "p_1"'] & A
\end{tikzcd}
$$

Now on the top left we have our pull back. However, on the top right, we've unraveled the pullback and ignored the terminal object to observe that $P$ has the universal property of what a product would demand. Hence we may denote $P = A \times B$ as the product. Thus by Proposition 3.3.8 $\mathcal{C}$ has all finite products. Note that we wouldn't have been able to construct this if we didn't have a terminal object; For example, if $\mathcal{C}$ was a discrete category, we wouldn't even have any morphisms to take a pullback on!

Now to derive equalizers, consider a pair of parallel morphisms $f, g : A \longrightarrow B$. Then we may simply take their pullback to obtain the diagram below.

$$
\begin{tikzcd}
C \arrow[dr, dashed, "h"] \arrow[drr, bend left, "f"] \arrow[ddr, bend right, "g"'] & & \\
& P \arrow[r, "p_1"] \arrow[d, "p_2"'] & A \arrow[d, "f"] \\
& A \arrow[r, "g"'] & B
\end{tikzcd}
$$

If $p : A \times A \longrightarrow A$ is the natural projection map, then because we have a trivial mapping $1_A : A \longrightarrow A$, there exists a canonical canonical map $i : A \longrightarrow A \times A$ such that $p \circ i = 1_A$. Similarly, because we have mappings $p_1, p_2 : P \longrightarrow B$, we must have a mapping $h : P \longrightarrow A \times A$.

$$
\begin{tikzcd}
& A \arrow[dl, "1_A"'] \arrow[d, dashed, "i"] \arrow[dr, "1_A"] & \\
A & A \times A \arrow[l, "p"] \arrow[r, "p"'] & A
\end{tikzcd}
\qquad
\begin{tikzcd}
& P \arrow[dl, "p_2"'] \arrow[d, dashed, "h"] \arrow[dr, "p_1"] & \\
A & A \times A \arrow[l, "p"] \arrow[r, "p"'] & A
\end{tikzcd}
$$

Now we can take the pullback on the morphism $h : P \longrightarrow A \times A$ and $i : A \longrightarrow A \times A$ to obtain the equalizer.

$$
\begin{tikzcd}
C \arrow[drr, bend left, "f"] \arrow[dr, dashed, "h"] \arrow[ddr, bend right, "g"'] & & \\
& E \arrow[r, "p_1"] \arrow[d, "p_2"'] & P \arrow[d, "h"] \\
& A \arrow[r, "i"'] & A \times A
\end{tikzcd}
$$

Hence we see that for finite limits, we can reduce our assumptions to pullbacks and a terminal object, giving rise to the final corollary.

**Theorem 5.3.4.** If a category has pullbacks and a terminal object, then it has all finite limits.

## 5.4 Preservation of Limits

**Definition 5.4.1.** Let $F : J \longrightarrow C$ be a diagram and suppose $G : \mathcal{C} \longrightarrow \mathcal{D}$ is a functor. If for every limit $\operatorname{Lim} F$ exists in $\mathcal{C}$ with morphisms $u_i : C \longrightarrow F_i$, we say $G$ **preserves limits** if $G(\operatorname{Lim} F)$ is a limit with morphisms $G(u_i) : G(C) \longrightarrow G(F_i)$. Moreover, we call such a functor a **continuous functor**.

As an immediate consequence of the definition, it should be noted that a composition of continuous functors is continuous.

Below we see a visual definition of a continuous functor.

$$
\begin{tikzcd}
 & C \arrow[d, dashed, "h"] \arrow[dl, bend right, "f_i"'] \arrow[dr, bend left, "f_j"] & \\
 & \operatorname{Lim} F \arrow[dl, "u_i"'] \arrow[dr, "u_j"] & \\
F(i) \arrow[rr, "F(g)"] & & F(j)
\end{tikzcd}
\qquad
\begin{tikzcd}
 & D \arrow[d, dashed, "h'"] \arrow[dl, bend right, "g_i"'] \arrow[dr, bend left, "g_j"] & \\
 & G(\operatorname{Lim} F) \arrow[dl, "G(u_i)"'] \arrow[dr, "G(u_j)"] & \\
G(F(i)) \arrow[rr, "G(F(g))"] & & G(F(j))
\end{tikzcd}
$$

There's one particular and important functor which is always continuous in any category.

**Theorem 5.4.2.** Let $\mathcal{C}$ be a small category. Then for each $C \in \mathcal{C}$, the functor

$$\operatorname{Hom}_{\mathcal{C}}(C, -) : \mathcal{C} \longrightarrow \mathbf{Set}$$

preserves limits. (Dually, the functor $\operatorname{Hom}_{\mathcal{C}}(-, C) = \operatorname{Hom}_{\mathcal{C}}(C, -) : \mathcal{C}^{\text{op}} \longrightarrow \mathbf{Set}$ takes colimits to limits.)

*Proof.* Let $F : J \longrightarrow \mathcal{C}$ be a diagram with a limiting object $\operatorname{Lim}\ F$ equipped with the morphisms $\sigma_i : \operatorname{Lim}\ F \longrightarrow F_i$. Then applying the $\operatorname{Hom}_{\mathcal{C}}(C, -)$ functor to $\operatorname{Lim}\ F$ and to each $u_i$, we realize it forms a cone in **Set**.

$$
\begin{tikzcd}
 & \operatorname{Lim}\ F \arrow[dl, "\sigma_i"'] \arrow[dr, "\sigma_j"] & \\
F_i \arrow[rr, "u"'] & & F_j
\end{tikzcd}
\qquad
\begin{tikzcd}
 & \operatorname{Hom}_{\mathcal{C}}(C, \operatorname{Lim}\ F) \arrow[dl, "\sigma_{i*}"'] \arrow[dr, "\sigma_{j*}"] & \\
\operatorname{Hom}_{\mathcal{C}}(C, F_i) \arrow[rr, "u_*"'] & & \operatorname{Hom}_{\mathcal{C}}(C, F_j)
\end{tikzcd}
$$

Now we show that $\operatorname{Hom}_{\mathcal{C}}(C, \operatorname{Lim}\ F)$, equipped with the morphisms $\sigma_{i*}$, is a universal cone; that is, it is a limit. Suppose that $X$ is a set which forms a cone with the morphisms $\tau_i : X \longrightarrow \operatorname{Hom}_{\mathcal{C}}(C, F_i)$.

$$\begin{tikzcd}
 & X \arrow[dl, "\tau_i"'] \arrow[dr, "\tau_j"] & \\
\text{Hom}_{\mathcal{C}}(C, F_i) \arrow[rr, "u_*"'] & & \text{Hom}_{\mathcal{C}}(C, F_j)
\end{tikzcd}$$

Then for each $x \in X$, we see that $\tau_i(x) : C \longrightarrow F_i$. The diagram above tells us that $u \circ \tau_i(x) = \tau_j(x)$ for each $x$. Hence each $x \in X$ induces a cone with apex $C$ with morphisms $\tau_i(x) : C \longrightarrow F_i$.

$$\begin{tikzcd}
 & C \arrow[dl, "\tau_i(x)"'] \arrow[dr, "\tau_j(x)"] & \\
F_i \arrow[rr, "u"'] & & F_j
\end{tikzcd}$$

However, Lim $F$ is the limit of $F : J \longrightarrow \mathcal{C}$. Therefore, there exists a unique arrow $h_x : C \longrightarrow \text{Lim } F$ such that $h_x \circ \sigma_i = \tau_i(x)$. Now we can uniquely define a function : $X \longrightarrow \text{Hom}_{\mathcal{C}}(C, \text{Lim } F)$ where $h(x) = h_x : C \longrightarrow \text{Lim } F$, in such a way that the diagram below commutes.

$$\begin{tikzcd}
 & X \arrow[ddl, bend right, "\tau_i"'] \arrow[d, dashed, "h"] \arrow[ddr, bend left, "\tau_j"] & \\
 & \text{Hom}_{\mathcal{C}}(C, \text{Lim } F) \arrow[dl, "\sigma_{i*}"'] \arrow[dr, "\sigma_{j*}"] & \\
\text{Hom}_{\mathcal{C}}(C, F_i) \arrow[rr, "u_*"'] & & \text{Hom}_{\mathcal{C}}(C, F_j)
\end{tikzcd}$$

Therefore, $\text{Hom}_{\mathcal{C}}(C, \text{Lim } F)$ is a limit in **Set**. ∎

At this point, you may be wondering: What is the difference between a functor which "creates limits" and one which preserves them? We'll see that their definitions are different, but creating limits is the same as preserving them

**Theorem 5.4.3.** Suppose $G : \mathcal{C} \longrightarrow \mathcal{D}$ creates limits for $F : J \longrightarrow \mathcal{C}$. If $G \circ F : J \longrightarrow \mathcal{D}$ has a limit in $\mathcal{D}$, then $G$ is continuous.

*Proof.* Suppose $F : J \longrightarrow \mathcal{C}$ has limit $\text{Lim}\, F$ in $\mathcal{C}$ with morphisms $v_i : \text{Lim}\, F \longrightarrow F_i$ for each $i \in J$. Further, suppose $G \circ F : J \longrightarrow \mathcal{D}$ has a limit $\text{Lim}\, G \circ F$ with morphisms $u_i : \text{Lim}\, G \circ F \longrightarrow G \circ F_i$.

Since $G : \mathcal{C} \longrightarrow \mathcal{D}$ creates limits, this implies the existence of a limiting object $X$ with morphisms $\sigma_i : X \longrightarrow F_i$ for $F : J \longrightarrow C$ where $G(X) = \operatorname{Lim} G \circ F$ and $G(\sigma_i) = u_i$. However, limiting objects are unique (by their universal properties). As they must be isomorphic, there exists an isomorphism $\varphi : X \longrightarrow \operatorname{Lim} F$ for which $v_i \circ \varphi = \sigma_i$. Thus we see that

$$G(\operatorname{Lim} F) \cong G(X) = \operatorname{Lim} G \circ F \qquad G(v_i \circ \varphi) = G(\sigma_i) = u_i.$$

Therefore, $G$ preserves limits and so is continuous. ∎

We have the following as a corollary.

**Corollary 5.4.4.** Suppose $G : \mathcal{C} \longrightarrow \mathcal{D}$ creates limits and $\mathcal{C}$ is complete. Then $\mathcal{D}$ is complete and $G$ preserves limits.

## 5.5 Adjoints on Limits

Consider the free monoid functor $F$ and the forgetful functor $U$, as below. Recall that they form an adjunction.

$$\mathbf{Set} \underset{U}{\overset{F}{\rightleftarrows}} \mathbf{Mon}$$

The way that we philosophically interpret this adjunction is as follows: For a set $X$, a monoid homomorphism $\varphi : F(X) \longrightarrow M$ gives rise to a unique set function $f : X \longrightarrow U(M)$. Conversely, a set function $g : X \longrightarrow U(M)$ gives rise to a unique monoid homomorphism $\psi : F(X) \longrightarrow M$.

We will now observe that these functors exhibit nice behavior.

- Recall that products in **Mon** are simply products of monoids, while products in **Set** are cartesian products. One can show that, for two monoids $M$, $N$, we have the isomorphism

  $$U(M \times N) \cong U(M) \times U(N).$$

  Regarding this functor's behavior, we say that the forgetful functor $U$ **preserves** products.

- We may ask if the converse holds: Does the free functor preserve products? The answer is no: Given two sets $X, Y$, it is generally not true that $F(X \times Y) \cong F(X) \times F(Y)$ (as monoids).

  An easy way to see this is to let $X = Y = \{\bullet\}$, the one point set. Then $F(\{\bullet\} \times \{\bullet\}) \cong F(\{\bullet\}) \cong \mathbb{Z}$, while $F(\{\bullet\}) \times F(\{\bullet\}) \cong \mathbb{Z} \times \mathbb{Z}$.

- What is interesting, however, is that the free functor *does* preserve coproducts. Recall that the coproduct in **Set** is the disjoint union, while the coproduct in **Mon** is the free product of monoids. Then it is true that, for two sets $X, Y$,

  $$F(X \amalg Y) \cong F(X) * F(Y).$$

Thus we see that we have two functors that separately preserve products and coproducts. This is actually very interesting; after all, a very useful question to ask about a functor is if it preserves products, coproducts, equalizers, etc. For example, the fundamental group functor preserves products, and this is an interesting result one usually proves a topology course.

We now explain why we have this nice behavior.

**Theorem 5.5.1.** Suppose $G : \mathcal{D} \longrightarrow \mathcal{C}$ is a right adjoint and $F : \mathcal{C} \longrightarrow \mathcal{D}$ is its left adjoint. Then $G$ preserves limits and $F$ preserves colimits.

Before a proof, we make some comments.

- An easy way to remember this is **RAPL**: "**R**ight **A**djoints **P**reserve **L**imits." (Speaking from experience, say it in your head a bunch of times or you'll forget.) If you can remember **RAPL**, then you can remember that, dually, left adjoints preserve colimits.
- The converse of this theorem does not hold.
- Typically, this proof is shown in one of two forms: It is "blackboxed" with a slick application of the Yoneda Lemma, which is not illuminating or useful for a new reader. Or, it is more usefully spelled out by showing that right adjoints preserve limits, and the second statement is obtained by "dualizing". For variety, we will show that left adjoints preserve colimits. Then, the reader can try proving themselves that right adjoints preserve limits.

*Proof.* Let $(\operatorname{Colim} H, \sigma_i : H(i) \longrightarrow \operatorname{Colim} H)$ be the colimit of the functor $H : J \longrightarrow \mathcal{C}$. This means that we have the universal diagram below.

$$
\begin{tikzcd}
 & A & \\
 & \operatorname{Colim} H \arrow[u, dashed] & \\
H(i) \arrow[ur, "\sigma_i"] \arrow[uur, bend left, "\gamma_i"] \arrow[rr, "H(f)"'] & & H(j) \arrow[ul, "\sigma_j"'] \arrow[uul, bend right, "\gamma_j"']
\end{tikzcd}
$$

Mapping this to $\mathcal{D}$ under $F : \mathcal{C} \longrightarrow \mathcal{D}$, we obtain the diagram

$$
\begin{tikzcd}
 & F(\operatorname{Colim} H) & \\
F(H(i)) \arrow[ur, "F(\sigma_i)"] \arrow[rr, "F(H(f))"'] & & F(H(j)) \arrow[ul, "F(\sigma_j)"']
\end{tikzcd}
$$

We see that $(F(\operatorname{Colim} H), F(\sigma_i) : F(H(i)) \longrightarrow F(\operatorname{Colim} H))$ is a cone over the functor $F \circ H : J \longrightarrow \mathcal{D}$. We must show it is universal. Towards that goal, let $(C, \tau_i : F(H(i)) \longrightarrow C)$ be a cone over $F \circ H : J \longrightarrow \mathcal{D}$. We must show that

1. There exists a $\alpha : F(\operatorname{Colim} H) \longrightarrow C$ such that $\alpha \circ F(\sigma_i) = \tau_i$ for all $i \in J$
2. $\alpha$ is the unique morphism from $F(\operatorname{Colim} H)$ to $C$ with this property.

We show existence. Observe that each $\tau_i : F(H(i)) \longrightarrow C$ induces a unique morphism $\delta_i : H(i) \longrightarrow G(C)$ such that the diagram below commutes.

$$
\begin{tikzcd}
F(G(C)) \arrow[r, "\varepsilon_C"] & C \\
F(H(i)) \arrow[u, dashed, "\delta_i"] \arrow[ur, "\tau_i"'] &
\end{tikzcd}
$$

Hence, we have a family of $\delta_i : H(i) \longrightarrow G(C)$. However, since $\operatorname{Colim} H$ is the colimit of $H$, we obtain a unique morphism $k : \operatorname{Colim} H \longrightarrow G(C)$ such that the diagram commutes.

$$\begin{tikzcd}
 & G(C) & \\
 & \operatorname{Colim} H \arrow[u, dashed, "k"] & \\
H(i) \arrow[uur, bend left, "\delta_i"] \arrow[ur, "\sigma_i"] \arrow[rr, "H(f)"'] & & H(j) \arrow[uul, bend right, "\delta_j"'] \arrow[ul, "\sigma_j"']
\end{tikzcd}$$

We then map this diagram in $\mathcal{C}$ to the diagram below in $\mathcal{D}$ via $F$:

$$\begin{tikzcd}
 & C & \\
 & F(G(C)) \arrow[u, "\varepsilon_C"] & \\
 & F(\operatorname{Colim} H) \arrow[u, "F(k)"] & \\
H(i) \arrow[uuur, bend left, "\tau_i"] \arrow[uur, bend left, "F(\delta_i)"] \arrow[ur, "F(\sigma_i)"] \arrow[rr, "F(H(f))"'] & & H(j) \arrow[uuul, bend right, "\tau_j"'] \arrow[uul, bend right, "F(\delta_j)"'] \arrow[ul, "F(\sigma_j)"']
\end{tikzcd}$$

Thus we see that $\varepsilon_C \circ F(k) : F(\operatorname{Colim} H) \longrightarrow C$ is a morphism pointing from $F(\operatorname{Colim})$ to $C$ such that the above diagram commutes. We have proved existence of such a morphism. It is not difficult to show uniqueness, which is left as an exercise. Once we have uniqueness, we can then conclude that $(F(\operatorname{Colim} H), F(\sigma_i) : F(H(i)) \longrightarrow F(\operatorname{Colim} H)$ forms a universal cone over $F \circ H : J \longrightarrow \mathcal{D}$, so that $F(\operatorname{Colim} H)$ is the colimit, as desired. ∎

---

**Example 5.5.2.** Using the above theorem, we now know that the free monoid functor $F : \mathbf{Set} \longrightarrow \mathbf{Mon}$ preserves coproducts. Therefore, we can say that for any sets $X, Y$, we have that

$$F(X \amalg Y) \cong F(X) * F(Y).$$

Moreover, the free monoid functor is part of a larger family of free functors:

- Free group functor, $F : \mathbf{Set} \longrightarrow \mathbf{Grp}$
- Free abelian group functor, $F : \mathbf{Set} \longrightarrow \mathbf{Ab}$

- Free ring functor, $F : \mathbf{Set} \longrightarrow \mathbf{Ring}$
- Free $R$-module functor, $F : \mathbf{Set} \longrightarrow R\text{-}\mathbf{Mod}$

who are the left adjoints to their respective forgetful functors. However, the coproduct in some of these categories is not always the free product. For example, the coproduct of **Grp** is the free product, but the coproduct in **Ab** is the direct sum. Hence, the above theorem tells us that coproducts are preserved, but to obtain the correct isomorphism, we need to remember what the coproduct in the codomain category of our left adjoint is.

---

**Example 5.5.3.** Let **Meas** be the category of measure spaces with measure-preserving morphisms. More precisely,

**Objects.** The objects are triples $(X, \mathcal{A}, \mu_X)$ where $X$ is a topological space, $\mathcal{A}$ is a sigma algebra on $X$, and $\mu_X$ is a measure on $X$.

**Morphisms.** A morphism between two objects $(X, \mathcal{A}, \mu_X)$ and $(Y, \mathcal{B}, \mu_Y)$ is a function $f : X \longrightarrow Y$ such that $f$ is measurable and preserves measure. That is, is $f$ is measurable and

$$\mu_X(f^{-1}(B)) = \mu_Y(B)$$

for every $B \in \mathcal{B}$.

Let $U : \mathbf{Meas} \longrightarrow \mathbf{Set}$ be the forgetful functor, forgetting measure space properties and measurability of the morphisms. This functor can't have a left-adjoint, since it does not preserve products. In fact, **Meas** cannot even have products. The main issue with this is that we cannot guarantee the projection morphisms to preserve measure. For example, if we consider the simple measure space $(\mathbb{R}, \mathcal{B}, \mu)$ where $\mathcal{B}$ consists of the Borel algebra and $\mu$ is the Lebesgue measure, then one reasonable way to try to form a product with itself is to construct the triple

$$(\mathbb{R} \times \mathbb{R}, \mathcal{B} \times \mathcal{B}, \mu \times \mu).$$

However, observe that the projection $\pi : (\mathbb{R} \times \mathbb{R}, \mathcal{B} \times \mathcal{B}, \mu \times \mu) \longrightarrow (\mathbb{R}, \mathcal{B}, \mu)$ is not measure preserving:

$$\mu \times \mu(\pi^{-1}([0, 1])) = \mu \times \mu([0, 1] \times \mathbb{R}) = \infty$$

while

$$\mu([0, 1]) = 0.$$

Therefore, we cannot form products. Hence our forgetful functor has no left adjoint.

One could guess that the left adjoint *would* be the measure-constructing functor $F : \mathbf{Set} \longrightarrow \mathbf{Meas}$ where

$$X \mapsto (X, \mathcal{P}, \mu_0)$$

where $\mathcal{P}$ is the sigma algebra on the power set, and $\mu_0$ assigns the measure of each set to zero (i.e. the trivial measure) but this is not the case. In fact, this functor itself also cannot have a left-adjoint because it doesn't preserve products (since **Meas** can't have products).

---

## Exercises

**1.** Denote the free monoid functor as $F$. Prove directly that for two sets $X$, $Y$, we have the isomorphism of monoids $F(X \amalg Y) \cong F(X) * F(Y)$. (Doing this is actually very important; The proof of Theorem 5.5.1 will become more intuitive.)

**2.** Finish the proof of Theorem 5.5.1

**3.** Let $\mathcal{C}, \mathcal{D}$ be categories with finite products.

*i.* Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor that preserves products, so that for two objects $A$, $B$ of $\mathcal{C}$, there exists an isomorphism

$$F(A \times B) \cong F(A) \times F(B).$$

Does this isomorphism have to be natural in $A, B$?

*ii.* Suppose $F : \mathcal{C} \longrightarrow \mathcal{D}$ is a right adjoint. Is the isomorphism $F(A \times B) \cong F(A) \times F(B)$ natural now?

## 5.6 Existence of Universal Morphisms and Adjoint Functors

When we introduced functors, we introduced several if and only if propositions which gave us criterion on the existence of an adjoint functor. Notably, we showed that if there exists an adjunction

$$C \underset{G}{\overset{F}{\leftrightarrows}} D$$

(that is, the classic bijection of homsets which is natural) then there exist universal morphisms

$$\eta_C : C \longrightarrow G \circ F(C) \qquad \varepsilon_D : F \circ G(D) \longrightarrow D$$

for all objects $C, D$. Furthermore, we only need one of the universal morphisms to derive an adjunction. Since universal morphisms are simply initial objects in some comma category, we have the following proposition.

**Proposition 5.6.1.** Let $G : \mathcal{D} \longrightarrow \mathcal{C}$ be a functor. Then $G$ has a left adjoint if and only if for each $C \in \mathcal{C}$, the comma category $C \downarrow G$ has an initial object.

*Proof.*

$\Longrightarrow$ Suppose $G$ has a left adjoint $F : \mathcal{C} \longrightarrow \mathcal{D}$. Then for each $C \in \mathcal{C}$, there exists a universal morphism $\eta_C : C \longrightarrow G(F(C))$. Now in the comma category, objects will be of the form

$$(D, f : C \longrightarrow G(D))$$

where morphisms between $(D, f : C \longrightarrow G(D))$ and $(D', f' : C \longrightarrow G(D'))$ will be induced by morphisms $h : D \longrightarrow D'$ such that

$$\begin{tikzcd}
& C \arrow[dl, "f"'] \arrow[dr, "f'"] & \\
G(D) \arrow[rr, "G(h)"'] & & G(D')
\end{tikzcd}$$

commutes. First, observe that $(F(C), \eta_C : C \longrightarrow G(F(C)))$ is an object of the comma category. Second, observe that the bijection of homsets

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

(natural in $C, D$) guarantees that every object $(D, f : C \longrightarrow G(D))$ in the comma category corresponds uniquely to a morphism $h : F(C) \longrightarrow D$. Moreover, uniqueness guarantees that the diagram

$$\begin{tikzcd} & C \arrow[dl, "\eta_C"'] \arrow[dr, "f"] & \\ G(F(C)) \arrow[rr, "G(h)"'] & & G(D) \end{tikzcd}$$

must commute. Hence, $(F(C), \eta_C : C \longrightarrow G(F(C)))$ is an initial object $C \downarrow G$.

$\Longleftarrow$ Now suppose that $C \downarrow G$ has an initial object $(D, \eta_C : C \longrightarrow G(D))$. Actually, denote the object $D$ as $F(C)$. When we write $F(C)$, we're not denoting a functor, because we'll show this is a functor. Anyways, our initial object can be written as

$$(F(C), \eta_C : C \longrightarrow G(F(C))).$$

This defines a mapping on objects $C \mapsto F(C)$. To show that this is a functor, suppose we have a morphism $f : C \longrightarrow C'$ in $\mathcal{C}$. Then we have square

$$\begin{tikzcd} C \arrow[r, "\eta_C"] \arrow[d, "f"'] & G(F(C)) \\ C' \arrow[r, "\eta_{C'}"] & G(F(C')). \end{tikzcd}$$

Adding the final leg to this diagram would show that $F$ is a functor. But since $(F(C), \eta_C : C \longrightarrow G(F(C)))$ is an initial object in $(C \downarrow G)$, and $(F(C'), \eta_{C'} : C' \longrightarrow G(F(C)))$ is an object in this category, there must be a *unique* morphism $F(f) : F(C) \longrightarrow F(C')$. Uniqueness of this morphism forces commutativity of the square

$$\begin{tikzcd} C \arrow[r, "\eta_C"] \arrow[d, "f"'] & G(F(C)) \arrow[d, "G(F(f))"] \\ C' \arrow[r, "\eta_{C'}"] & G(F(C')). \end{tikzcd}$$

and therefore $F$ is a functor. Simultaneously, this shows $F$ is left adjoint to $G$, as desired.

∎

We can repeat the proof to achieve the following result as well.

**Corollary 5.6.2.** Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a functor. Then $F$ has a right adjoint if and only if for each $D \in D$, the comma category $D \downarrow F$ has a terminal object.

Thus we see that initial and terminal objects are key to figuring out when a functor has a left or right adjoint, and hence when they preserve limits. We can investigate a little deeper into this.

**Lemma 5.6.3.** (Initial Object Existence.) If $\mathcal{C}$ is a complete category with small homsets, then $\mathcal{C}$ has an initial object if and only if it satisfies the **Solution Set Condition**:

> There exists objects $(C_i)_{i\in I} \in \mathcal{C}$ such that for every $C \in \mathcal{C}$, there is a a morphism $f_i : C_i \longrightarrow C$ for at least one $i \in I$.

*Proof.*

$\Longrightarrow$ Suppose $\mathcal{C}$ has an initial object $C'$. Then $I$ is the one-point set since for each $C \in C$ there exists one unique morphism $f : C' \longrightarrow C$.

$\Longleftarrow$ On the other hand, assume the solution set condition. Since $\mathcal{C}$ is complete, it must have products, so we may take the product

$$W = \prod_{i\in J} C_i.$$

This product has associated projection morphisms $\pi_k : \prod_{i\in J} C_i \longrightarrow C_k$. Therefore, for each object $C \in \mathcal{C}$, there exists at least one morphism between $W$ and $C$ by composition:

$$f_k \circ \pi_k : W \longrightarrow C.$$

By hypothesis, the collection of endomorphisms $\text{Hom}_{\mathcal{C}}(W, W)$ is a set. Therefore, we may form an equalizer $e : V \longrightarrow W$ of this set. Observe that for each $C \in \mathcal{C}$, there exists at least one morphism between $V$ and $C$ by composition:

$$f_k \circ \pi_k \circ e : V \longrightarrow C.$$

We'll now show that all morphisms are equal. Suppose the contrary; that there are two distinct morphisms $f, g : V \longrightarrow C$. Denote the equalizer of this pair as $e_1 : u \longrightarrow v$. Then we have that

$$\begin{tikzcd}
U \arrow[r, "e_1"] & V \arrow[r, "f"] \arrow[d, "e"] & C \\
W \arrow[u, dashed, "s"] \arrow[r, "e\circ e_1\circ s"'] & W = \prod_{i\in J} C_i \arrow[r, "\pi_k"'] & C_k \arrow[u, "f_i"']
\end{tikzcd}$$

commutes. The morphism $s$ is induced via the universality of both $U$ and $V$. Since $e \circ e_1 \circ s : W \longrightarrow W$, and $e$ is the equalizer of endomorphisms of $W$, we have that

$$(e \circ e_1 \circ s) \circ e = e.$$

Since equalizers are monic, we can cancel on the left side to conclude that

$$e_1 \circ s \circ e = 1_V.$$

However, this implies that the right inverse of $e_1$ is $s \circ e$. Since $e_1$ is already monic, it must be an isomorphism. Hence $f = g$, so that $V$ is an initial object as desired.

∎

We can now combine all of our propositions and theorems into the following one, which is the main adjoint functor theorem of interest.

**Theorem 5.6.4. (General Adjoint Functor Theorem.)** Let $\mathcal{D}$ be complete with small homsets. A functor $G : \mathcal{D} \longrightarrow \mathcal{C}$ has a left adjoint if and only if it preserves all small limits and satisfies the **solution set condition**:

For each $C \in \mathcal{C}$, there exists a set of objects $(D_i)_{i \in I}$ $\mathcal{D}$ and a family of arrows

$$f_i : C \longrightarrow G(D_i)$$

such that for every morphism $h : C \longrightarrow G(D)$, there exists a $j \in I$ and a morphism $t : D_j \longrightarrow D$ such that

$$h = G(t) \circ f_i.$$

The above theorem helps us find out when we can get a left adjoint. Prior to this theorem, we already know what happened if we were given a functor who has a left adjoint. Namely, it must preserve limits. This natural question one would then ask is if the converse holds. The above theorem tells us no, the converse doesn't hold and in fact we need to make sure the functor satisfies the **solution set condition**. In the next section, we'll give an example of a functor which preserves limits from a complete category, but still has no left adjoint.

As a converse to the above theorem, we have the following.

**Theorem 5.6.5. (Representability Theorem.)** Let $\mathcal{C}$ be a small, complete category. A functor $K : \mathcal{C} \longrightarrow \textbf{Set}$ is representable if and only if $K$ preserves limits and satisfies the following **solution set condition**:

There exists a set $S \subseteq \text{Ob}(\mathcal{C})$ such that for any $C \in \mathcal{C}$ and any $x \in K(C)$, there exists an $s \in S$, an element $y \in K(s)$ and an arrow

$$f : s \longrightarrow C \text{ such that } K(f)(y) = x.$$

## 5.7 Subobjects and Quotient Objects

The entire point of category theory, contrary to its name, is to unify mathematics. Mathematicians saw the same stories over and over again in algebra and topology, and one day they got sick of it and decided to start naming the patterns they were seeing. Mathematicians achieved a level of abstraction where we no longer really care about the objects, but we want to study the morphisms between them. However, in many categories, the objects are often things like groups, rings, or topological spaces; hence there are subgroups, subrings, and spaces with subset topologies which also exist inside categories we study. This presents a challenge for category theory: how do we generalize the notion of subgroups or subspaces if we always avoid explicit reference to the elements?

It turns out that the correct way to go about this is to consider the philosophy of sub-"things": whenever $S$ is a sub-"thing" of $X$, there usually exists a monomorphism

$$m : S \longrightarrow X.$$

For example, in **Set**, $S \subseteq X$ implies that there's an injection $i : S \longrightarrow X$; a monomorphism is injective in **Set**, so this makes sense. In **Top**, if $S \subseteq X$ where $S$ is given the subspace topology, then the inclusion function $i : S \longrightarrow X$ is continuous, so there does exist a monomorphism $m : S \longrightarrow X$ in **Top**.

Thus we see that these monomorphisms give us sub-"things," and so we might naively say the set of all "subobjects" of an object $X$ in a category $\mathcal{C}$ is the set

$$\text{Sub}_{\mathcal{C}}(X) = \{S \in \text{Ob}(\mathcal{C}) \mid \exists f : S \longrightarrow X \text{ with } f \text{ monic } \}.$$

However, the space of all of these monomorphisms is huge, and also repetitive. For example, in **Set**, if we have $X = \{1, 2, 3, 4, 5\}$, then there are all kinds of monomorphisms into $X$:

$$\begin{tikzcd}
\{!\ ,\ \%\ ,\ \$\ ,\&\ \} \arrow[dr, hook] & \\
\{Q,G,X,I\} \arrow[r, hook] & \{1,2,3,4,5\} \\
\{\text{All}\ ,\ \text{Politicians}\ ,\ \text{Are}\ ,\ \text{Corrupt}\ \} \arrow[ur, hook] &
\end{tikzcd}$$

Each arrow is basically saying the same thing. How do we deal with this? Well, we can impose an equivalence relation on this space to obtain something smaller and more manageable.

Let $A$ an object of our category $\mathcal{C}$. Consider monomorphisms $f : C \longrightarrow A$ and $g : D \longrightarrow A$. Define the relation $\leq$ on monomorphisms of this form where

$f \leq g$ if there exists an $h$ where $f = g \circ h$.

$$\begin{tikzcd} C \arrow[dr, hook, "f"] & \\ & A \\ D \arrow[ur, hook, "g"'] \arrow[uu, dashed, "h"] & \end{tikzcd}$$

for some monomorphism $h : D' \longrightarrow D$. Note that if $f \leq g$ and $f \geq g$, then $C$ and $D$ are isomorphic (this is not true in general; this only true here because $f, g$ are monomorphisms). So we now have our equivalence relation: we say $f \sim g$ if there exists an isomorphism $\varphi : D \longrightarrow C$ which makes the above diagram commute.

**Definition 5.7.1.** Let $\mathcal{C}$ be a category and let $A$ be an object. We say a **subobject** of $A$ is an equivalence class of monomorphisms $f : S \longrightarrow A$ under the equivalence relation $\sim$. We denote this space of equivalence classes as

$$\text{Sub}_{\mathcal{C}}(A) = \Big\{[f] \mid f : C \longrightarrow A \text{ is a monomorphism}\Big\}.$$

---

**Example 5.7.2.** Let $\mathcal{C}$ be a category. An interesting application of subobjects occurs in functor categories. To illustrate this we consider the functor category $\mathbf{Set}^{\mathcal{C}}$; that is, the category with functors $F : \mathcal{C} \longrightarrow \mathbf{Set}$ whose morphisms are natural transformation $\eta : F \longrightarrow G$ between such functors.

If we play around with these functors long enough, we may ask the question: What happens when, for a functor $F : \mathcal{C} \longrightarrow \mathbf{Set}$, there is another functor $G : \mathcal{C} \longrightarrow \mathbf{Set}$ such that

$$G(A) \subseteq F(A)?$$

Could we logically call $G$ a "**subfunctor**" of $F$? We could with a little more work. Because $G(A) \subseteq F(A)$, we know that there exists a monomorphism (just an injection here) $i_A : G(A) \longrightarrow F(A)$. Now a natural question to ask here is if this translates to a natural transformation. That is, does the diagram below commute?

$$\begin{tikzcd} A \arrow[d, "f"] & G(A) \arrow[r, hook, "i_A"] \arrow[d, "G(f)"'] & F(A) \arrow[d, "F(f)"] \\ B & G(B) \arrow[r, hook, "i_B"'] & F(B) \end{tikzcd}$$

The answer is no. This is because $G(f)$ and $F(f)$ could be two entirely different functions which do two entirely different things to the same elements in different domains; however, one way for this diagram to commute is if $G(f)$ is $F(f)$ *restricted* to the set $G(A)$. That is, if

$$G(f) = F(f)\Big|_{G(A)}.$$

The diagram then commutes. But is this the only way to make it commute? Suppose with no assumption of $G(f)$ that the diagram did commute. Then we can still make a morphism $F(f)\big|_{G(A)} : G(A) \longrightarrow G(B)$ to get the commutative diagram

$$
\begin{tikzcd}
A \arrow[d, "f"] & G(A) \arrow[d, shift right, "F(f)|_{G(A)}"'] \arrow[d, shift left, "G(f)"] \arrow[r, hook, "i_A"] & F(A) \arrow[d, "F(f)"] \\
B & G(B) \arrow[r, hook, "i_B"'] & F(B)
\end{tikzcd}
$$

Then we see that $i_B \circ G(f) = i_B \circ F(f)\big|_{G(A)}$. However, $i_B$ is a monomorphism, so $G(f) = F(f)\big|_{G(A)}$. Hence the *only* way to make the diagram commute is if $G(f)$ is a restriction of $F(f)$.

Thus we could define $G : \mathcal{C} \longrightarrow \mathbf{Set}$ to be a subfunctor of $F : \mathcal{C} \longrightarrow \mathbf{Set}$ if $G(A) \subseteq F(A)$ and $G(f : A \longrightarrow B) = F(f)\big|_{G(A)}$. Or, equivalently, if $G(A) \subseteq F(A)$ and that this relation is natural.

However, we can recover the same concept by applying subobjects to this functor category. In this case, we can (with laziness) say a $G : \mathcal{C} \longrightarrow \mathbf{Set}$ is a subobject of the functor $F : \mathcal{C} \longrightarrow \mathbf{Set}$ in $\mathbf{Set}^{\mathcal{C}}$ if there exists a monic natural transformation $\eta : G \longrightarrow F$.

Unwrapping this definition, we see that a monic natural transformation in this case is just one where each morphism $\eta_A : G(A) \longrightarrow F(A)$ is a monomorphism, which, in our case, just means an inclusion function, such that the necessary square commutes. However, we already showed that we get the commutativity of the necessary square if and only if $G(f : A \longrightarrow B) = F(f)\big|_{G(A)}$.

Hence we have recovered the same concept of a **subfunctor** in two different ones; one in which we followed our intuition, and one in which we blinded applied the concept of a subobject in the functor category $\mathbf{Set}^{\mathcal{C}}$.

---

The previous example allows us to make the definition:

**Definition 5.7.3.** Let $\mathcal{C}, \mathcal{D}$ be categories. Then a functor $G : \mathcal{C} \longrightarrow \mathcal{D}$ is a **subfunctor** of $F : \mathcal{D} \longrightarrow \mathcal{C}$ if $G$ is a subobject of $F$ in the functor category $\mathcal{D}^{\mathcal{C}}$.

Now, perhaps unsurprisingly, the entire process above can be dualized. When we dualize, however, we obtain a generalization of the concept of quotient objects. Instead of just dualizing and being boring, we'll motivate why we'd even care for such a dual concept.

In interesting categories such as **Ab** or **Top**, we not only have subgroups and subspaces, but we also have quotient groups and quotient spaces. For the case of abelian groups, we can, for any such group $G$, consider any subgroup $H \leq G$ and construct the quotient group $G/H$. This comes with a a nice epimorphism $\pi : G \longrightarrow G/H$ where $g \mapsto g + H$.

For topological spaces $(X, \tau)$ in **Top**, we can define an equivalence relation $\sim$ on $X$ and consider the topological space $(X/\sim, \tau')$ such that $\tau'$ is the topology where a set $U$ is open if $\{x \mid [x] \in U\}$ is open in $\tau$. We can then equip ourselves with a continuous projection map $\pi : X \longrightarrow X/\sim$, which is also an epimorphism.

With these few examples, we see that it is worthwhile to generalize the concept of quotient objects; to do this however requires no explicit mention of the elements of the objects of the category. However, we can maintain the philosophy seen in the previous two examples to generalize the concept.

For an object $A$ in a category $\mathcal{C}$, we consider all *epimorphisms*

$$e : A \longrightarrow Q$$

and call objects such objects $Q$ as quotient objects. Again, the space of these objects is too large, so we instead consider ordering relation

$f \leq g$ if there exists an $h$ where $f = h \circ g$.

$$\begin{tikzcd} & C \\ A \arrow[ur, two heads, "f"] \arrow[dr, two heads, "g"'] & \\ & D \arrow[uu, dashed, "h"'] \end{tikzcd}$$

Observing that $f \leq g$ and $g \leq f$ together imply that $C \cong D$, we see that we may construct an equivalence relation $\sim$ where $f \sim g$ if there exists an isomorphism $\varphi : D \longrightarrow C$ such that $f = \varphi \circ g$. We can now outline a clear definition.

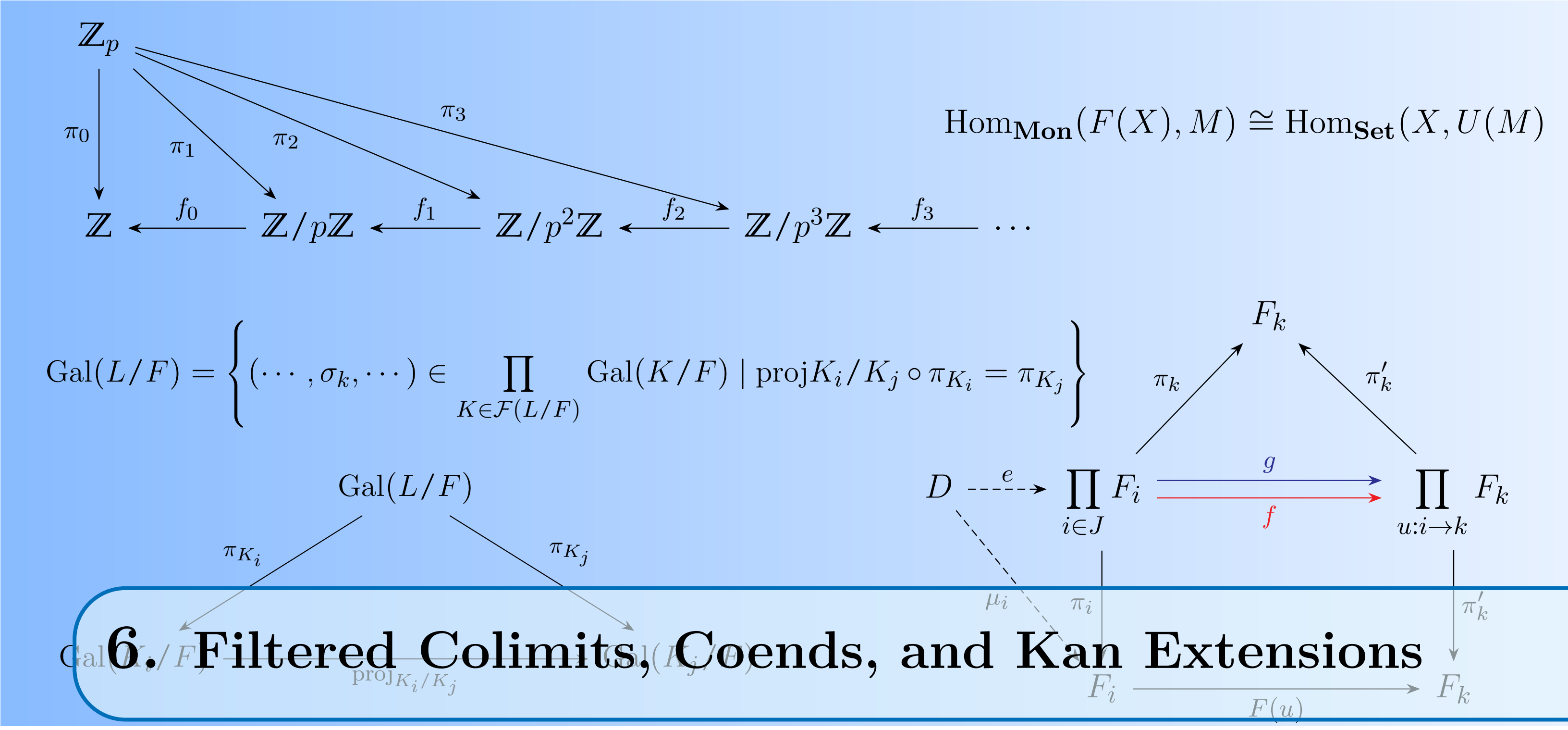
**Definition 5.7.4.** Let $\mathcal{C}$ be a category and let $A$ be an object. We say a **quotient object** of $A$ is an equivalence class of morphisms $f : A \longrightarrow Q$. We then denote

$$\text{Quot}_{\mathcal{C}}(A) = \Big\{[f] \mid f : A \longrightarrow Q \text{ is an epimorphism } \Big\}.$$

---

**Example 5.7.5.** A quotient object in **Cat** is a quotient category (from chapter 2)

---

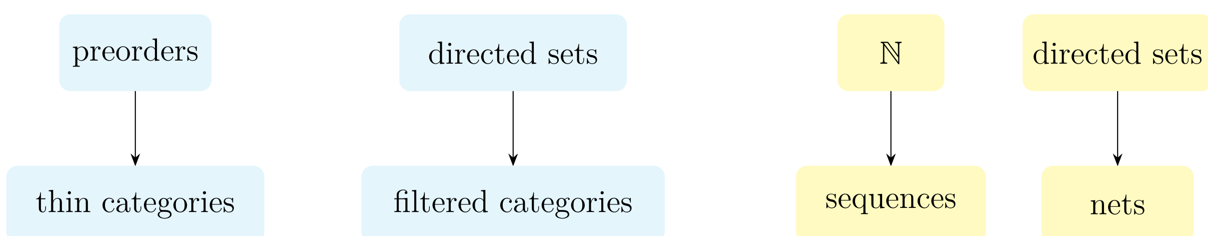


# 6. Filtered Colimits, Coends, and Kan Extensions

## 6.1 Filtered Categories and Limits

Outside of category theory, the most common types of limits that are taken in areas such as algebraic geometry and topology are inverse and directed limits. These are limits which are taken over thin categories (or preorders) which have at most one morphism between any two morphisms.

As we shall see, limits over thin categories do not possess the nice properties that limits taken over *filtered* categories have, which we will see is the categorification of the notion of a *directed set*. We will motivate our desire to work with filtered categories instead of just thin categories by observing an analogous motivation to work with directed sets instead of $\mathbb{N}$ in sequences within topology. The picture in mind should be:




*On the left, we see that thin and filtered categories are the categorification of concepts which we will use to take limits over. On the right, we have topology concepts of sequences and nets, which are limits taken over different sets.*

Let $X$ be a topological space. Recall that a sequence $\{a_n\}_{n=1}^{\infty}$ in $X$ is a function $a : \mathbb{N} \longrightarrow X$ such that $a(n) = a_n$. We say the sequence converges to a point $x \in X$ if for every open set $U$ of $x$ there exists a $N \in \mathbb{N}$ such that $\{a_N, a_{N+1}, \dots,\} \subseteq U$.

Some of the first topological spaces that people worked with were metric spaces $(X, d)$, and the properties of these spaces were worked out over time. People eventually figured out that

- A subset $F \subseteq X$ is closed if and only if $F$ contains the limits of every sequence in $F$.
- A subset $U \subseteq X$ is open if and only if $U$ contains does not contain the limit of any sequence in $X - U$.

This is a wonderful result! However, it does not generalize to arbitrary topological spaces. There are weird counterexamples that we will not get into (cite An Introduction to Topology and Homotopy Theory by Sierdaski).

What this means is that sequences over a plain preorder (i.e., $\mathbb{N}$) are great, and they have nice properties, but they lack the ability to extend their nice properties to arbitrary topological spaces. We need more if we want it to work over arbitrary spaces.

This is where a directed set comes in.

**Definition 6.1.1.** A **directed set** $D$ is a set equipped with a binary relation $\le$ such that for all $a, b, c \in D$,

**1.** $a \le a$ (Reflexive).
**2.** if $a \le b$ and $b \le c$, then $a \le c$ (Transitive)
**3.** For all $a, b \in D$, there exists a $c \in C$ such that $a \le c$ **and** $b \le c$ (Directed).

The first two properties describe a preorder; only the last condition is new to us. To summarize, the "directed" axiom grants us an upper bounded in $D$ for any finite set of elements of $D$.

Let $D$ be a directed set. Define a **net**, or **Moore-Smtih Sequence**, to be a function $\lambda : D \longrightarrow X$. We say a net $\lambda$ converges to a point $x \in X$ if for every open set $U$ containing $x$, there exists a $d \in D$ such that $\{\lambda(c) \mid c \ge d\} \subseteq U$.

Directed sets are then enough to give us the following theorem:

**Theorem 6.1.2.** Let $X$ be a topological space.

- A subset $F \subseteq X$ is closed if and only if every convergent net $\lambda : D \longrightarrow X$ has a limit in $F$
- A subset $U \subseteq X$ is open if and only if every convergent net $\lambda : D \longrightarrow X - U$ does not have a limit in $U$.

Hence we see that limits taken over preorders have substantial benefits than when they are simply taken over $\mathbb{N}$. Similarly, what we will see is that limits taken over filtered categories enjoy much better properties than limits simply taken over preorders. First, we introduce filtered categories.

**Definition 6.1.3.** We say that a category $J$ is **filtered** if

1. For any pair of objects $j, j'$, there exists an object $k$ and morphism $u : j \longrightarrow k$ and $v : j' \longrightarrow k$.
2. For any pair of parallel morphism $u, v : i \longrightarrow j$, there exists an object $k$ and a morphism $w : j \longrightarrow k$ such that the diagram below commutes.

We do not say the empty category is filtered; this should be obvious, but it also needs to be said.


*Conditions (1) and (2) illustrated.*

---

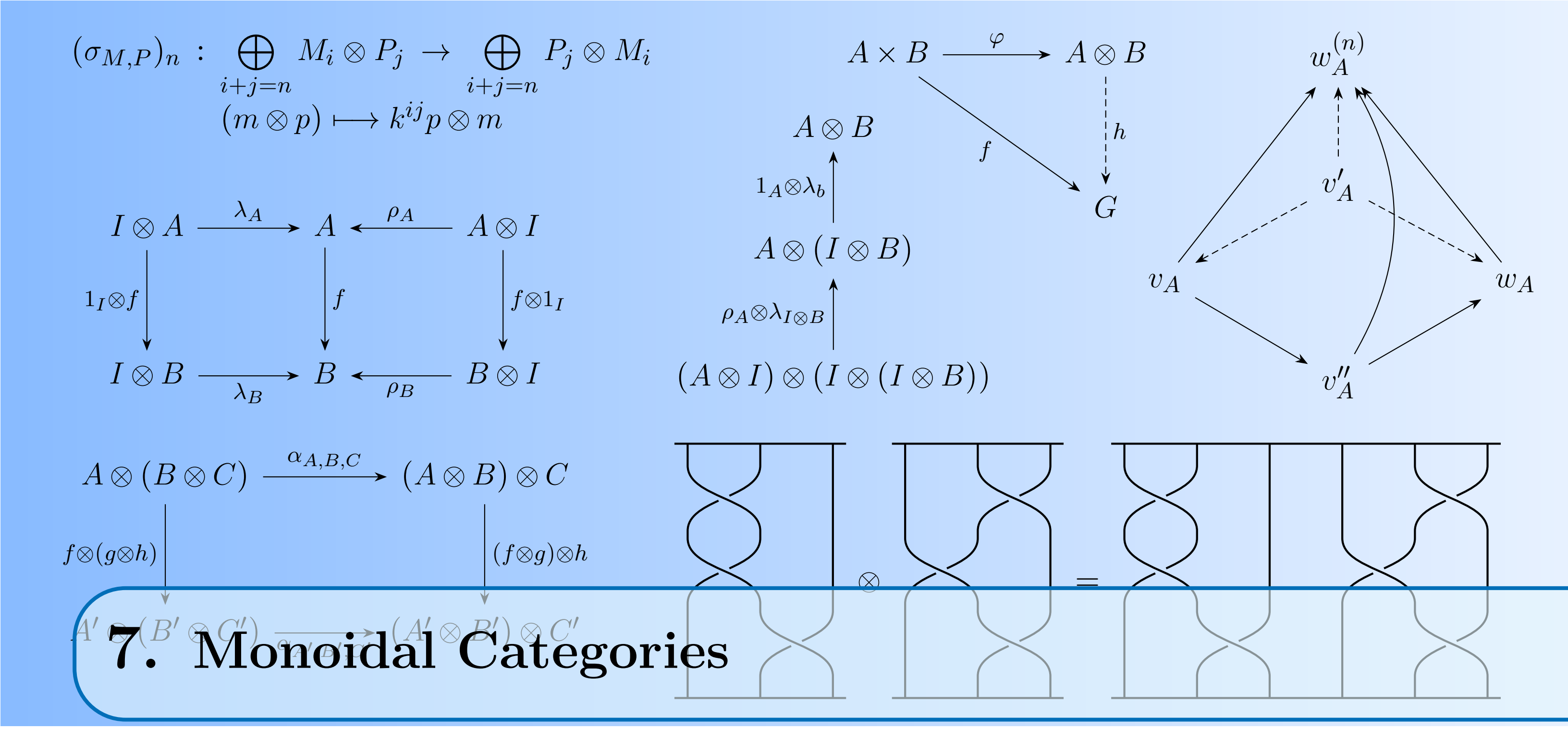
**Example 6.1.4.** Let $J$ be a thin category. What does it take for $J$ to be filtered? Well, in a thin category, there is never any pair of distinct morphisms. Hence condition (2) is trivial. Therefore, for $J$ to be filtered, we simply need to satisfy (1). But in the language of thin categories, condition (1) can be read as "for any $j, j \in J$, there exists a $k$ such that $j, j' \leq k$". Such a condition holds if and only if

$$\text{every finite subset } S \subseteq J \text{ has an upper bound } \textit{in } J.$$

Thus, a thin category $J$ needs to have the above property in order to be a filtered category.

An example of this concerns the category $\mathbf{Open}(X)$, where $X$ is a topological space. The objects are open sets, while morphisms are inclusions. The maximal element $X \in \mathbf{Open}(X)$ always exists, and hence makes this thin category filtered.

---



# 7. Monoidal Categories

## 7.1 Monoidal Categories

The concept of a monoidal category is motivated by the very simple observation that some categories are canonically equipped with their own algebraic data which allows us to multiply objects of the category to get new objects. This is similar to how in a group $G$, we multiply two group elements $g$, $h$ to get another group element $gh \in G$. These types of categories appear frequently enough in many settings that it has been necessary to really understand what the core ingredients of these categories are. The task of defining these categories, however, takes a bit of work. Before we offer the definition and discuss such work we motivate monoidal categories with two key examples.

---

**Example 7.1.1.** Consider the category **Set**. Then for two sets $A$, $B$, we can take their cartesian product to create a third set

$$A \times B = \{(a, b) \mid a \in A, b \in B\}.$$

We also know that given three sets $A, B, C$, we have an isomorphism $A \times (B \times C) \cong (A \times B) \times C$. The bijection is given by the function

$$\alpha_{A,B,C} : A \times (B \times C) \xrightarrow{\sim} (A \times B) \times C \qquad (a, (b, c)) \mapsto ((a, b), c).$$

In addition, there is a particularly special set $\{\bullet\}$, the one element set. For this set, we know that $\{\bullet\} \times A \cong A \times \{\bullet\} \cong A$. The bijections are given by

$$\lambda_A : \{\bullet\} \times A \xrightarrow{\sim} A \qquad (\bullet, a) \mapsto a$$
$$\rho_A : A \times \{\bullet\} \xrightarrow{\sim} A \qquad (a, \bullet) \mapsto a$$

A final observation that is easy to check is that our morphisms $\alpha_{A,B,C}$, $\lambda_A$, and $\rho_A$ are natural. Naturality for $\alpha$ means that for any three functions $f : A \longrightarrow A'$, $g : B \longrightarrow B'$, $h : C \longrightarrow C'$, the diagram below commutes

$$\begin{tikzcd}
A \times (B \times C) \arrow[r, "\alpha_{A,B,C}"] \arrow[d, "f\times(g\times h)"'] & (A \times B) \times C \arrow[d, "(f\times g)\times h"] \\
A' \times (B' \times C') \arrow[r, "\alpha_{A',B',C'}"'] & (A' \times B') \times C'
\end{tikzcd}$$

while naturality for $\lambda$ and $\rho$ means that for any function $f : A \longrightarrow A'$, the diagrams below commute.

$$\begin{tikzcd}
\{\bullet\} \times A \arrow[r, "\lambda_A"] \arrow[d, "1\times f"'] & A \arrow[d, "f"] \\
\{\bullet\} \times A' \arrow[r, "\lambda_{A'}"'] & A'
\end{tikzcd}
\qquad
\begin{tikzcd}
A \times \{\bullet\} \arrow[r, "\rho_A"] \arrow[d, "f\times 1"'] & A \arrow[d, "f"] \\
A' \times \{\bullet\} \arrow[r, "\rho_{A'}"'] & A'
\end{tikzcd}$$

(Here, 1 denotes the identity $1 : \{\bullet\} \longrightarrow \{\bullet\}$). While being able to find the functions $\alpha, \rho, \lambda$ and observing that they are natural may not be surprising in **Set**, what is surpising is that this behavior continues in many other categories.

---

---

**Example 7.1.2.** Let $k$ be a field, and consider the category $\mathbf{Vect}_k$ of vector spaces over $k$. For two vector spaces $U$, $V$, we may take their tensor product to create a third vector space over $k$. There are many ways to describe $U \otimes V$; here is one of them:

$$U \otimes V = \left\{ u \otimes v \;\middle|\; u \in U,\, v \in V \quad \text{such that} \quad \begin{array}{l} 1.\ (u_1 + u_2) \otimes v = u_1 \otimes v + u_2 \otimes v \\ 2.\ u \otimes (v_1 + v_2) = u \otimes v_1 + u \otimes v_2 \\ 3.\ c(u \otimes v) = (cu) \otimes v = u \otimes (cv), c \in k \end{array} \right\}$$

Moreover, if $U$, $V$ have bases $\{e_i\}_{i\in I}$, $\{f_j\}_{j\in J}$, then the basis of $U \otimes V$ is $\{e_i \otimes f_j\}_{(i,j)\in I\times J}$.

From linear algebra, we know that $U \otimes (V \otimes W) \cong (U \otimes V) \otimes W$. To show this, we will define an isomorphic linear transformation $U \otimes (V \otimes W) \longrightarrow (U \otimes V) \otimes W$. However, recall that to define such a linear transformation, it suffices to define it on the basis. Thus, let $W$ have basis $\{g_\ell\}_{\ell\in L}$. Then we define

$$\alpha_{U,V,W} : U \otimes (V \otimes W) \xrightarrow{\sim} (U \otimes V) \otimes W$$

where on the basis elements

$$\alpha_{U,V,W}(e_i \otimes (f_j \otimes g_\ell)) = (e_i \otimes f_j) \otimes g_\ell.$$

This establishes our desired isomorphism.

In addition, the field $k$ is trivially a vector space over itself; its basis is the multiplicative identity 1. Moreover, we have the isomorphisms $k \otimes V \cong V \otimes k \cong V$. The isomorphisms are given by the linear transformations

$$\lambda_V : k \otimes V \xrightarrow{\sim} V \qquad 1 \otimes e_i \mapsto e_i$$
$$\rho_V : V \otimes k \xrightarrow{\sim} V \qquad e_i \otimes 1 \mapsto e_i$$

Here, we've defined the two transformations on the bases.

Similarly to our last example, we comment that $\alpha, \lambda, \rho$ defined here are natural. This means that for any three linear transformations $f : U \longrightarrow U'$, $g : V \longrightarrow V'$, and $h : W \longrightarrow W'$, the diagram below commutes

$$\begin{tikzcd}
U \otimes (V \otimes W) \arrow[r, "\alpha_{U,V,W}"] \arrow[d, "f\otimes(g\otimes h)"'] & (U \otimes V) \otimes W \arrow[d, "(f\otimes g)\otimes h"] \\
U' \otimes (V' \otimes W') \arrow[r, "\alpha_{U',V',W'}"'] & (U' \otimes V') \otimes W'
\end{tikzcd}$$

and we additionally have that the diagrams below commute.

$$\begin{tikzcd}
k \otimes U \arrow[r, "\lambda_U"] \arrow[d, "1\otimes f"'] & U \arrow[d, "f"] \\
k \otimes U' \arrow[r, "\lambda_{U'}"'] & U'
\end{tikzcd}
\qquad
\begin{tikzcd}
U \otimes k \arrow[r, "\rho_U"] \arrow[d, "f\otimes 1"'] & U \arrow[d, "f"] \\
U' \otimes k \arrow[r, "\rho_{U'}"'] & U'
\end{tikzcd}$$

---

The observations we have made here continue to be true upon investigating many other categories $\mathcal{C}$ in which we have some known, natural way of combining elements. In each case, the story is the same. The key ingredients are:

- There is some product $\otimes : \mathcal{C} \times \mathcal{C} \longrightarrow \mathcal{C}$ (specifically, it is a bifunctor)
- For all $A, B, C \in \mathcal{C}$, there is a natural isomorphism

$$\alpha_{A,B,C} : A \otimes (B \otimes C) \xrightarrow{\sim} (A \otimes B) \otimes C$$

- There is a special object $I$ of $\mathcal{C}$ such that, for any object $A$, we have the natural isomorphisms

$$\lambda_A : I \otimes A \xrightarrow{\sim} A \qquad \rho_A : A \otimes I \xrightarrow{\sim} A$$

The fact that we keep seeing these patterns in many categories is what motivates the following definition.

**Definition 7.1.3.** A **monoidal category** $\mathcal{C} = (\mathcal{C}, \otimes, I)$ is a category $\mathcal{C}$ equipped with a bifunctor $\otimes : \mathcal{C} \times \mathcal{C} \longrightarrow \mathcal{C}$, a (special) object $I$, and three natural isomorphisms

$$\begin{aligned}
\alpha_{A,B,C} &: A \otimes (B \otimes C) \xrightarrow{\sim} (A \otimes B) \otimes C && \textbf{(Associator)}\\
\lambda_A &: I \otimes A \xrightarrow{\sim} A && \textbf{(Left Unit)}\\
\rho_A &: A \otimes I \xrightarrow{\sim} A && \textbf{(Right Unit)}
\end{aligned}$$

such that the following **coherence conditions** hold.

$$\begin{tikzcd}
A \otimes (I \otimes B) \arrow[rr, "\alpha_{A,I,B}"] \arrow[dr, "1_A \otimes \lambda_B"'] & & (A \otimes I) \otimes B \arrow[dl, "\rho_A \otimes 1_B"] \\
& A \otimes B &
\end{tikzcd} \tag{7.1}$$

$$\begin{tikzcd}
A \otimes (B \otimes (C \otimes D)) \arrow[r, "\alpha_{A,B,C\otimes D}"] \arrow[d, "1_A \otimes \alpha_{B,C,D}"'] & (A \otimes B) \otimes (C \otimes D) \arrow[r, "\alpha_{A\otimes B,C,D}"] & ((A \otimes B) \otimes C) \otimes D \\
A \otimes ((B \otimes C) \otimes D) \arrow[rr, "\alpha_{A,B\otimes C,D}"'] & & (A \otimes (B \otimes C)) \otimes D \arrow[u, "\alpha_{A,B,C} \otimes 1_D"']
\end{tikzcd} \tag{7.2}$$

We also define some terminology within this definition.

- We call the bifunctor $\otimes$ the **monoidal product**
- We refer to $I$ as the **identity object**
- We refer to diagram 7.1 as the **unit diagram** and diagram 7.2 as the **pentagon diagram**.

Further, we say a **strict monoidal category** is one in which the associator, left unit and right unit are all identities.

The reader should be wondering: What are those "coherence conditions"? The short answer is that we need the coherence conditions in order for our ideas to make any logical sense. While that answer is very vague and unsatisfying, we are not quite yet ready to fully explain why those two diagrams are necessary. We will however say

- The reader is definitely not expected at this moment to understand why those diagrams are necessary.
- We will eventually explain why those diagrams are necessary.

Before we explain why the diagrams are necessary, we develop further intuition regarding monoidal categories with some more examples.

---

**Example 7.1.4.** As one might expect, $(\mathbf{Set}, \times, \{\bullet\})$ is a monoidal category. We have verified most of the details except the coherence conditions, but it is not too difficult to show that the unit and pentagon diagram commute in **Set**.

However, we can put another monoidal category structure on **Set** with the following data:

- We let the disjoint union bifunctor $(-)\coprod(-) : \mathbf{Set}\times\mathbf{Set}\longrightarrow\mathbf{Set}$ be our monoidal product.
- We let the empty set $\varnothing$ be our identity object.

With these settings, we can define natural isomorphisms for any three sets $X, Y, Z$

- $\alpha_{X,Y,Z} : X \amalg (Y \amalg Z) \xrightarrow{\sim} (X \amalg Y) \amalg Z$
- $\lambda_X : \varnothing \amalg X \xrightarrow{\sim} X$
- $\rho_X : X \amalg \varnothing \xrightarrow{\sim} X$

in the obvious way, and check that the required diagrams commute. In this way, we have that $(\mathbf{Set}, \amalg, \varnothing)$ is also a monoidal category.

---

The previous example demonstrates that a given category can have more than one monoidal structure on it. This is analagous to the fact that sometimes one can put two different group structures on an underlying set.

The previous example may also make us wonder if we can generalize our logic to consider other categories in which finite products and coproducts exist. The answer is yes, and this gives us many examples of monoidal categories:

- $(\mathbf{Top}, \times, \{\bullet\})$
- $(\mathbf{Ab}, \oplus, \{e\})$
- $(\text{R-}\mathbf{Mod}, \times, \{0\})$

**Proposition 7.1.5.** If $\mathcal{C}$ is a category with finite products and a terminal object $T$, then $(\mathcal{C}, \times, T)$ is a monoiodal category. We refer to this type of monoidal category as a **cartesian monoidal category**.

Dually, if it has finite coproducts and an initial object $I$, then $(\mathcal{C}, \amalg, I)$ is also a monoidal category. We call this type of monoidal category a **cocartesian monoidal category**.

We now introduce less obvious, but useful examples of monoidal categories.

---

**Example 7.1.6.** Let $R$ be a commutative ring. Then the category of all $R$-modules, ($R$-**Mod**, $\otimes$, $\{0\}$), forms a monoidal category under the tensor product. Recall that the tensor product between two $R$-modules $N \otimes M$ is an initial object in the comma category $(N \times M \downarrow R\text{-}\mathbf{Mod}$ where the morphisms are bilinear. Alternatively in diagrams, we have that

$$\begin{tikzcd} M \times N \arrow[r, "\varphi"] \arrow[dr, "f"'] & M \otimes N \arrow[d, dashed, "h"] \\ & K \end{tikzcd}$$

Now consider a third $R$-module $P$; then we have two ways of constructing the tensor product. To demonstrate that we may identify these objects up to isomorphism, construct the maps

$$f : (M \otimes N) \times P \longrightarrow M \otimes (N \otimes P) \qquad \left(\sum_i m_i \otimes n_i, p\right) \mapsto \sum_i m_i \otimes (n_i \otimes p)$$

and

$$f' : M \times (N \otimes P) \longrightarrow (M \otimes M) \otimes P \qquad \left(m, \sum_j n_j \otimes p_j\right) \mapsto \sum_j (m \otimes n_j) \otimes p_j.$$

These maps are bilinear due to the bilinearity of $\otimes$. Hence we see that the universal property of the tensor product gives us unique map $\alpha$ and $\alpha'$ such that the diagrams below commute.

$$\begin{tikzcd} (M \otimes N) \times P \arrow[r, "\varphi"] \arrow[dr, "f"'] & (M \otimes N) \otimes P \arrow[d, dashed, "\alpha"] \\ & M \otimes (N \otimes P) \end{tikzcd} \qquad \begin{tikzcd} M \otimes (N \times P) \arrow[r, "\varphi'"] \arrow[dr, "f'"'] & M \otimes (N \otimes P) \arrow[d, dashed, "\alpha'"] \\ & (M \otimes N) \otimes P \end{tikzcd}$$

Based on how we defined $f$ and $f'$, and since we know that $\varphi$ and $\varphi'$ is, we can determine that $\alpha$ and $\alpha'$ are "shift maps", i.e,

$$\alpha\left(\sum_i (m_i \otimes n_i) \otimes p_i\right) = \sum_i m_i \otimes (n_i \otimes p_i) \qquad \alpha'\left(\sum_i m_i \otimes (n_i \otimes p_i)\right) = \sum_i (m_i \otimes n_i) \otimes p_i.$$

Hence we see that $\alpha$ and $\alpha'$ are inverses, so what we have is an associator:

$$\alpha_{M,N,P} : (M \otimes N) \otimes P \xrightarrow{\sim} M \otimes (N \otimes P).$$

Now consider the trivial $R$-module, denoted $I = \{0\}$. For any $R$-module $M$ we have evident maps

$$\sum_i 0 \otimes m_i \mapsto m_i \qquad \sum_i m_i \otimes 0 \mapsto 0$$

which provide isomorphisms, so that we have left and right associators

$$\lambda_M : I \otimes M \xrightarrow{\sim} M \qquad \rho_M : M \otimes I \xrightarrow{\sim} M.$$

Finally, the triangular and pentagonal diagrams are commutative since shifting the tensor product on individual elements does not change (up to isomorphism) the value of the overall elements.

---

---

**Example 7.1.7.** Consider the category $\mathbf{GrMod}_R$ which consist of graded $R$-modules $M = \{M_n\}_{n=1}^{\infty}$ Then this forms a monoidal category $(\mathbf{GrMod}_R, \otimes, I)$ where $I = \{(0)_n\}_{n=1}^{\infty}$ is the trivial graded $R$-module and where we define the monoidal product as $M \otimes N = \{(M \otimes N)_n\}_{n=1}^{\infty}$ where

$$(M \otimes N)_n = \bigoplus_{i+j=n} M_i \otimes N_j.$$

To show this monoidal, the first thing we must check is that we have an associator. Towards this goal, consider three graded $R$-modules $M = \{M_n\}_{n=1}^{\infty}$, $N = \{N_n\}_{n=1}^{\infty}$ and $P = \{P_n\}_{n=1}^{\infty}$. Then the $m$-th graded module of $M \otimes (N \otimes P)$ is

$$\begin{aligned}
[M \otimes (N \otimes P)]_m = \bigoplus_{i+j=m} M_i \otimes (N \otimes P)_j &= \bigoplus_{i+j=m} M_i \otimes \left( \bigoplus_{h+k=j} N_h \otimes P_k \right) \\
&= \bigoplus_{i+h+k=m} M_i \otimes (N_h \otimes P_k) \\
&\cong \bigoplus_{i+h+k=m} (M_i \otimes N_h) \otimes P_k \\
&= \bigoplus_{l+k=m} \left( \bigoplus_{i+h=l} M_i \otimes N_h \right) \otimes P_k \\
&= \bigoplus_{l+k=m} (M \otimes N)_l \otimes P_k \\
&= [M \otimes (N \otimes P)]_m
\end{aligned}$$

where in the third step we used the fact that the tensor product commutes with direct sums and in the fourth step we used the canonical associator regarding the tensor products of three element. Thus we see that we have an associator

$$\alpha : M \otimes (N \otimes P) \xrightarrow{\sim} (M \otimes N) \otimes P$$

which as a graded module homomorphism, acts on each level as

$$\alpha_m : [M \otimes (N \otimes P)]_m \xrightarrow{\sim} [(M \otimes N) \otimes P]_m$$

where in each coordinate of the direct sums we apply an instance of the associator $\alpha'$ between the tensor product of three $R$-modules. The naturality of this associator is inherited from $\alpha'$. In addition, we have natural left and right unitors

$$\lambda_M : I \otimes M \xrightarrow{\sim} M \qquad \rho_M : M \otimes I \xrightarrow{\sim} M$$

where on each level we utilize the natural left and right unitors for non-graded $R$-modules.

---

---

**Example 7.1.8.** Let $(M, \otimes, I, \alpha, \rho, \lambda)$ be a monoidal category, $\mathcal{C}$ any other category. Then the functor category $\mathcal{C}^M$ is a monoidal category. We treat the constant functor $I : \mathcal{C} \longrightarrow M$ where

$$I(A) = I \text{ for all } A$$

as the identity element, and we can define a tensor product on this category as follows: on objects $F, G : \mathcal{C} \longrightarrow M$, we define $F \boxtimes G$ as the composite

$$F \boxtimes G : \mathcal{C} \longrightarrow \xrightarrow{\Delta} \mathcal{C} \times \mathcal{C} \xrightarrow{(F \times G)} M \times M \xrightarrow{\otimes} M$$

which can be stated pointwise as $(F \boxtimes G)(C) = F(C) \otimes G(C)$. On morphisms, we have that if $\eta : F_1 \longrightarrow F_2$ and $\eta' : G_1 \longrightarrow G_2$ are natural transformations, then we say $\eta \boxtimes \eta' : F_1 \boxtimes G_1 \longrightarrow F_2 \boxtimes G_2$ is a natural transformation, where we define

$$(\eta \boxtimes \eta')_A = \eta_A \otimes \eta'_A : F_1(A) \otimes G_1(A) \longrightarrow F_2(A) \otimes G_2(A).$$

Note that such a natural transformation is well-defined as the diagram below commutes

$$\begin{tikzcd}
A \arrow[d, "f"] & F_1(A) \otimes G_1(A) \arrow[r, "\eta_A \otimes \eta'_A"] \arrow[d, "F_1(f) \otimes G_1(f)"'] & F_2(A) \otimes G_2(A) \arrow[d, "F_2(f) \otimes G_2(f)"] \\
B & F_1(B) \otimes G_1(B) \arrow[r, "\eta_B \otimes \eta'_B"'] & F_2(A) \otimes G_2(A)
\end{tikzcd}$$

since $\otimes : M \times M \longrightarrow M$ is a bifunctor. Finally, for functors $F, G, H : \mathcal{C} \longrightarrow M$ define the associator $\alpha'_{F,G,H} : F \boxtimes (G \boxtimes H) \xrightarrow{\sim} (F \boxtimes (G \boxtimes H))$ as the natural transformation where for each object $A$

$$(\alpha'_{F,G,H})_A = \alpha_{F(A),G(A),H(A)} : F(A) \otimes (G(A) \otimes H(A)) \longrightarrow (F(A) \otimes G(A)) \otimes H(A)$$

and the unitors $\lambda'_F : I \boxtimes F \longrightarrow F$ and $\rho'_F : F \boxtimes I \longrightarrow F$ as the natural transformations where for each object $A$

$$(\lambda'_F)_A = \lambda_A : I \otimes F(A) \longrightarrow F(A) \qquad (\rho'_F)_A = \rho_A : F(A) \otimes I \longrightarrow F(A).$$

One can then show that these together satisfy the pentagon and unit axioms.

---

---

**Example 7.1.9.** Consider the category $\mathbb{P}$ whose objects are the natural numbers (with 0

included) and whose morphisms are the symmetric groups $S_n$. That is,
**Objects.** The objects are $n = 0, 1, 2, \dots$.
**Morphisms.** For any objects $n, m$ we have that

$$\mathrm{Hom}_{\mathbb{P}}(n,m) = \begin{cases} S_n & \text{if } n = m \\ \varnothing & \text{if} n \neq m. \end{cases}$$

Note that there are many ways of constructing this category; we just present the simplest. In general terms this is the countable disjoint union of the symmetric groups. Even more generally, this can be done for any family of groups (or rings, monoids, semigroups).

What is interesting about this category is that it intuitively forms a strict monoidal category. That is, we can formulate a bifunctor $+ : \mathbb{P} \times \mathbb{P} \longrightarrow \mathbb{P}$ on objects as addition of natural numbers and on morphisms as

$$\sigma \otimes \tau \in S_{n+m}$$

where $\sigma \in S_n$ and $\tau \in S_m$ and where $\sigma \otimes \tau$ denotes the **direct sum permutation**. I could tell you in esoteric language and notation what that is, or I could just show you: $\sigma$ and $\tau$, displayed as below

$$\begin{array}{ccc} (1, 2, \dots, n) & & (1, 2, \dots, m) \\ \big\downarrow & & \big\downarrow \\ (\sigma(1), \sigma(2), \dots, \sigma(n)) & & (\tau(1), \tau(2), \dots, \tau(m)) \end{array}$$

become $\sigma \otimes \tau$ which is displayed as below.

$$\begin{array}{c} (1, 2, \dots, n, n+1, n+2, \dots, n+m) \\ \big\downarrow \\ (\sigma(1), \sigma(2), \dots, \sigma(n), n+\tau(1), n+\tau(2), \dots, n+\tau(m)) \end{array}$$

To make this monoidal, we specify that 0 is our identity element whose associated identity morphism is the empty permutation. Now clearly this operation is strict on objects. On morphisms, it is also strict in the same way that stacking three Lego pieces together in the two possible different ways are equivalent. Hence the associators and unitors are all identities and this forms a strict monoidal category.

---

**Proposition 7.1.10.** Let $\mathcal{C}$ be a category equipped with natural isomorphisms

$$\begin{aligned} \alpha_{A,B,C} &: A \otimes (B \otimes C) \xrightarrow{\sim} (A \otimes B) \otimes C \\ \lambda_A &: I \otimes A \xrightarrow{\sim} A \\ \rho_A &: A \otimes I \xrightarrow{\sim} A \end{aligned}$$

for all objects $A, B, C \in \mathcal{C}$ and some identity object $I$. Suppose the pentagonal diagram 7.2 holds for all objects of $\mathcal{C}$. Then for all $A, B \in \mathcal{C}$, the diagram

$$
\begin{tikzcd}
A \otimes (I \otimes B) \arrow[rr, "\alpha_{A,I,B}"] \arrow[dr, "1_A \otimes \lambda_B"'] & & (A \otimes I) \otimes B \arrow[dl, "\rho_A \otimes 1_B"] \\
& A \otimes B &
\end{tikzcd}
$$

commutes if and only if the diagram

$$
\begin{tikzcd}
A \otimes (B \otimes I) \arrow[rr, "\alpha_{A,B,I}"] \arrow[dr, "1_A \otimes \rho_B"'] & & (A \otimes B) \otimes I \arrow[dl, "\rho_{A \otimes B}"] \\
& A \otimes B &
\end{tikzcd}
$$

commutes, which commutes if and only if the diagram

$$
\begin{tikzcd}
I \otimes (A \otimes B) \arrow[rr, "\alpha_{I,A,B}"] \arrow[dr, "\lambda_{A \otimes B}"'] & & (I \otimes A) \otimes B \arrow[dl, "\rho_A \otimes 1_B"] \\
& A \otimes B &
\end{tikzcd}
$$

commutes.

This ultimately tells us that the definition of a monoidal category is not unique. That is, there are two different yet exactly equivalent ways in which we could have defined a monoidal category.

Thus what we see is that the definition of a monoidal category is very vast, and asks for a lot. Moreover, we've shown that there are different coherence conditions we could have imposed (the pentagonal diagram being the same), but they amount to stating the same thing.

## 7.2 Monoidal Functors

**Definition 7.2.1.** Let $(\mathcal{C}, \otimes, I)$ and $(\mathcal{D}, \odot, J)$ be monoidal categories. A **(lax) monoidal functor** is a functor $F : \mathcal{C} \longrightarrow \mathcal{D}$ equipped with the following data.

- For each pair $A$, $B$ in $\mathcal{C}$, we have a natural morphism

$$\varphi_{A,B} : F(A) \odot F(B) \longrightarrow F(A \otimes B)$$

  such that for any third object $C$, the diagram below commutes. (Note that we suppress the subscripts for clarity.)

$$\begin{tikzcd}
F(A) \odot \Big(F(B) \odot F(C)\Big) \arrow[r, "\alpha"] \arrow[d, "1\odot\varphi"'] & \Big(F(A) \odot F(B)\Big) \odot F(C) \arrow[d, "\varphi\odot 1"] \\
F(A) \odot F(B \otimes C) \arrow[d, "\varphi"'] & F(A \otimes B) \odot F(C) \arrow[d, "\varphi"] \\
F\Big(A \otimes \Big(B \otimes C\Big)\Big) \arrow[r, "F(\alpha)"'] & F\Big(\Big(A \otimes B\Big) \otimes C\Big)
\end{tikzcd}$$

- There exists a unique morphism $\varepsilon : J \longrightarrow F(I)$ such that, for any object $A$ of $\mathcal{C}$, the diagrams below commute. (Again, we suppress the subscripts for clarity.)

$$\begin{tikzcd}
F(A) \odot J \arrow[r, "\rho"] \arrow[d, "1\odot\varepsilon"'] & F(A) \\
F(A) \odot F(I) \arrow[r, "\varphi"'] & F(A \otimes I) \arrow[u, "F(\rho)"']
\end{tikzcd}
\qquad
\begin{tikzcd}
J \odot F(A) \arrow[r, "\lambda"] \arrow[d, "\varepsilon\odot 1"'] & F(A) \\
F(J) \odot F(A) \arrow[r, "\varphi"'] & F(I \otimes A) \arrow[u, "F(\lambda)"']
\end{tikzcd}$$

We say the $F$ is **strong** if $\varphi$ and $\varepsilon$ are isomorphisms and **strict** if $\varphi$ and $\varepsilon$ are identities.

We also define a **monoidal natural transformation** between two monoidal functors $\eta : F \longrightarrow G$ to be a natural transformation between the functors such that, for every $A, B$, the diagram below commutes.

$$\begin{tikzcd}
F(A) \odot F(B) \arrow[r, "\varphi_F"] \arrow[d, "\eta_A\odot\eta_B"'] & F(A \otimes B) \arrow[d, "\eta_{A\otimes B}"] \\
G(A) \odot G(B) \arrow[r, "\varphi_G"'] & G(A \otimes B)
\end{tikzcd}$$

**Example 7.2.2.** Consider the power set functor $\mathcal{P} : \mathbf{Set} \longrightarrow \mathbf{Set}$ which associates each set $X$ with its power set $\mathcal{P}(X)$. We may ask if this yields a monoidal functor

$$\mathcal{P} : (\mathbf{Set}, \times, \{\bullet\}) \longrightarrow (\mathbf{Set}, \times, \{\bullet\})$$

in any sense of lax, strong, or strict. It turns out that we may define a lax monoidal functor, but not a strong or strict.

Towards defining a lax monoidal functor, let $A, B$ two sets. Define $\varphi_{A,B} : \mathcal{P}(A) \times \mathcal{P}(B) \longrightarrow \mathcal{P}(A \times B)$ to be a function where if $U, V$ are subsets of $A, B$ respectively, then

$$\varphi_{A,B}(U, V) = U \times V.$$

In addition, we define the function $\varepsilon : \{\bullet\} \longrightarrow P(\{\bullet\})$ where

$$\varepsilon(\bullet) = \{\bullet\}.$$

Observe that with this data we have that for any sets $A, B, C$, the diagram below commutes

$$\begin{tikzcd}
\mathcal{P}(A) \times \Big(\mathcal{P}(B) \times \mathcal{P}(C)\Big) \arrow[r, "\alpha"] \arrow[d, "1\times\varphi"'] & \Big(\mathcal{P}(A) \times \mathcal{P}(B)\Big) \times \mathcal{P}(C) \arrow[d, "\varphi\times 1"] \\
\mathcal{P}(A) \times \mathcal{P}(B \times C) \arrow[d, "\varphi"'] & \mathcal{P}(A \times B) \times \mathcal{P}(C) \arrow[d, "\varphi"] \\
\mathcal{P}\Big(A \times \Big(B \times C\Big)\Big) \arrow[r, "\mathcal{P}(\alpha)"'] & \mathcal{P}\Big(\Big(A \times B\Big) \times C\Big)
\end{tikzcd}$$

and that for any set $A$ the diagrams below commute.

$$\begin{tikzcd}
\mathcal{P}(A) \times \{\bullet\} \arrow[r, "\rho"] \arrow[d, "1\times\varepsilon"'] & \mathcal{P}(A) \\
\mathcal{P}(A) \times \mathcal{P}(\{\bullet\}) \arrow[r, "\varphi"'] & \mathcal{P}(A \times \{\bullet\}) \arrow[u, "\mathcal{P}(\rho)"']
\end{tikzcd}
\qquad
\begin{tikzcd}
\{\bullet\} \times \mathcal{P}(A) \arrow[r, "\lambda"] \arrow[d, "\varepsilon\times 1"'] & \mathcal{P}(A) \\
\mathcal{P}(\{\bullet\}) \times \mathcal{P}(A) \arrow[r, "\varphi"'] & \mathcal{P}(\{\bullet\} \times A) \arrow[u, "\mathcal{P}(\lambda)"']
\end{tikzcd}$$

Note that our choice that $\varepsilon(\bullet) = \{\bullet\}$ was necessary in order for the above two diagrams to commute.

We now show that this cannot be a strong or strict monoidal functor. To see this, let $A, B$ be two sets. If $|X|$ denotes the cardinality of a set $X$, then observe that

$$|\mathcal{P}(A) \times \mathcal{P}(B)| = 2^{|A|} \cdot 2^{|B|} = 2^{|A|+|B|}$$

while
$$|\mathcal{P}(A \times B)| = 2^{|A\times B|}.$$

We see that in general these two sets are not of the same cardinality, and therefore one cannot establish an isomorphism between these two sets for all $A, B$, which we would need to do to at least construct a strong monoidal functor. Hence, we cannot regard this functor as strong or strict monoidal.

---

---

**Example 7.2.3.** The category of pointed topological spaces $\mathbf{Top}^*$ is the category where

**Objects.** Pairs $(X, x_0)$ with $X$ a topological space and $x_0 \in X$

**Morphisms.** A morphism $f : (X, x_0) \longrightarrow (Y, y_0)$ is given by a continuous function $f : X \longrightarrow Y$ such that $f(x_0) = y_0$.

This category is what allows us to characterize the fundamental group of a topological space as a functor
$$\pi_1 : \mathbf{Top}^* \longrightarrow \mathbf{Grp}$$
which sends a pointed space $(X, x_0)$ to its fundamental group $\pi_1(X, x_0)$ with $x_0$ as the selected basepoint. We demonstrate that this can be regarded as a monoidal functor
$$\pi_1 : \Big(\mathbf{Top}^*, \times, (\{\bullet\}, \bullet)\Big) \longrightarrow (\mathbf{Grp}, \times, \{e\})$$
where $\{e\}$ is the trivial group. The reader may be wondering how we are putting a cartesian product structure on the $\mathbf{Top}^*$, so we explain: For two topological spaces $X, Y$, we define
$$(X, x_0) \times (Y, y_0) = (X \times Y, (x_0, y_0))$$
where $X \times Y$ is given the product topology. The identity object $(\{\bullet\}, \bullet)$ is the trivial topological space with basepoint $\bullet$.

For any two pointed topological spaces $(X, x_0), (Y, y_0)$, define the function $\varphi_{X,Y} : \pi_1(X, x_0) \times \pi_1(Y, y_0) \longrightarrow \pi(X \times Y, (x_0, y_0))$ where for two loops $\beta, \gamma$ based as $x_0, y_0$ respectively, then
$$\varphi_{X,Y}(\beta, \gamma) = \beta \times \gamma : [0, 1] \longrightarrow X \times Y$$
which is in fact a loop in $X \times Y$ based at $(x_0, y_0)$. The above function is bijective; an inverse can be constructed by sending a loop $\delta$ in $X \times Y$ based at $(x_0, y_0)$ to the tuple $(p \circ \delta, q \circ \delta)$ where
$$p : X \times Y \longrightarrow X \qquad q : X \times Y \longrightarrow Y$$
are the continuous projection maps. It is not difficult to see that this preserves group products, so that $\varphi_{X,Y}$ establishes the isomorphism
$$\pi_1(X \times Y, (x_0, y_0)) \cong \pi_1(X, x_0) \times \pi_1(Y, y_0)$$

a fact usually proved in a topological course. In addition, this isomorphism to be natural: for two pointed topological spaces $(X, x_0)$ and $(Y, y_0)$, and for a pair of base-point preserving continuous functions $f : (X, x_0) \longrightarrow (W, w_0)$ and $g : (Y, y_0) \longrightarrow (Z, z_0)$, the following diagram commutes.

$$\begin{tikzcd} \pi_1(X, x_0) \times \pi_1(Y, y_0) \arrow[r, "\varphi_{X,Y}"] \arrow[d, "\pi_1(f)\times\pi_1(g)"'] & \pi_1(X \times Y, (x_0, y_0)) \arrow[d, "\pi_1(f\times g)"] \\ \pi_1(W, w_0) \times \pi_1(Z, z_0) \arrow[r, "\varphi_{W,Z}"'] & \pi_1(W \times Z, (w_0, z_0)) \end{tikzcd}$$

Thus $\varphi_{X,Y}$ is our desired natural isomorphism.

Next, define $\varepsilon : \{e\} \longrightarrow \pi_1(\{\bullet\}, \bullet)$ to be the group homomorphism that takes $e$ to the trivial loop at $\bullet$. As in the previous example, we are actually forced to define $\varepsilon$ in this way since $\{e\}$ is initial in **Grp**.

With this data, one can easily check that the necessary diagrams are commutative, so that the fundamental group functor $\pi_1$ is strong monoidal.

---

**Example 7.2.4.** Recall that a **Lie algebra** is a vector space $\mathfrak{g}$ over a field $k$ with a bilinear function $[-,-] : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}$ such that

**Antisymmetry.** For all $x, y \in \mathfrak{g}$, $[x, y] = -[y, x]$

**Jacobi Identity.** For all $x, y, z \in \mathfrak{g}$ we have that

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0.$$

For every Lie algebra $\mathfrak{g}$, we may create the **universal enveloping algebra** $U(\mathfrak{g})$. This is the algebra constructed as follows: If $T(\mathfrak{g})$ is the tensor algebra of $\mathfrak{g}$, i.e.,

$$T(\mathfrak{g}) = k \oplus (\mathfrak{g} \otimes \mathfrak{g}) \oplus (\mathfrak{g} \otimes \mathfrak{g} \otimes \mathfrak{g}) \oplus \cdots$$

and $I(\mathfrak{g})$ is the ideal generated by elements of the form $x \otimes y - y \otimes x - [x, y]$, then

$$U(\mathfrak{g}) = T(\mathfrak{g})/I(\mathfrak{g}).$$

By Corollary V.2.2(b) of [**?**], this construction is actually a functor

$$U : \mathbf{LieAlg} \longrightarrow k\text{-}\mathbf{Alg}.$$

Both categories can be regarded monoidal: $(\mathbf{LieAlg}, \oplus, \{\bullet\})$ is the monoidal category where we apply the cartesian product between Lie algebras, and $(k\text{-}\mathbf{Alg}, \otimes, k)$ is the monoidal category where we apply tensor products between $k$-algebras over the field $k$. The associators and unitors are the same that we have encountered in previous examples of monoidal categories with cartesian and tensor products.

We demonstrate that the universal enveloping algebra functor is strong monoidal:

$$U : (\mathbf{LieAlg}, \oplus, \{\bullet\}) \longrightarrow (k\text{-}\mathbf{Alg}, \otimes, k)$$

- By Corollary V.2.3 of [?], we have that if $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are two Lie algebras then $U(\mathfrak{g}_1 \oplus \mathfrak{g}_2) \cong U(\mathfrak{g}_1) \otimes U(\mathfrak{g}_2)$. One can use Corollary V.2.3(a) to show that this isomorphism is natural in both $\mathfrak{g}_1$ and $\mathfrak{g}_2$. We let this morphism be our required isomorphism

$$\varphi_{\mathfrak{g}_1,\mathfrak{g}_2} : U(\mathfrak{g}_1 \oplus \mathfrak{g}_2) \longrightarrow U(\mathfrak{g}_1) \otimes U(\mathfrak{g}_2).$$

- Note that $U(\{\bullet\}) = k$. Therefore, we let $\varepsilon : k \longrightarrow k$ be the identity.

As the associators and unitors are simple for monoidal categories with cartesian and tensor products, it is not difficult to show that the required diagrams commute. In this case, what is more difficult is obtaining naturality in $\varphi$, although this is taken care of (in a long proof) in Kassel's text.

## 7.3 What are those Coherence Conditions?

We are now going to address the elephant in the room: we have not explained why we have included diagrams 7.1 and 7.2 in our definition. To explain this, we are going to discuss the general structure of a monoidal category and answer the natural questions that arise.

Let $(\mathcal{M}, \otimes, I, \alpha, \rho, \lambda)$ be a monoidal category. For objects $A, B, C, \dots$ of $\mathcal{M}$, we can use the monoidal product $\otimes$ to generate various new expressions such as $A \otimes B$ which represent different objects in $\mathcal{M}$. Observe that using three objects, there are two different ways to combine the 3 objects:

$$A \otimes (B \otimes C) \qquad (A \otimes B) \otimes C.$$

There are five ways to combine 4 objects:

$$A \otimes (B \otimes (C \otimes D)) \quad A \otimes ((B \otimes C) \otimes D) \quad ((A \otimes B) \otimes C) \otimes D$$
$$A \otimes ((B \otimes C) \otimes D) \quad (A \otimes (B \otimes C)) \otimes D.$$

And there are 14 ways to combine 5 objects. We will not list them here.

Initially, we don't really know what the relationship is between the various expressions we are generating. For example, we may naturally wonder if

$$A \otimes (B \otimes C) \text{ and } (A \otimes B) \otimes C$$

or

$$A \otimes (B \otimes (C \otimes D)) \text{ and } A \otimes ((B \otimes C) \otimes D)$$

have *any relation* with each other. This is because in practice when $A, B, C, D$ are sets, vector spaces, groups, or whatnot, the above expressions *do* have something to do with each other. As we have seen, that relationship is usually an isomorphism. Therefore, if we are to develop some kind of theory of monoidal categories which we can apply to real mathematics, we ought to make sure that these objects are isomorphic in some way.

Monoidal categories by definition do in fact provide isomorphisms between different choices of multiplying together a set of objects. For example, from the axioms of a monoidal category, we know that the objects $A \otimes (B \otimes C)$ and $(A \otimes B) \otimes C$ are related via the natural isomorphism $\alpha_{A,B,C}$.

$$A \otimes (B \otimes C) \xrightarrow{\alpha_{A,B,C}} (A \otimes B) \otimes C$$

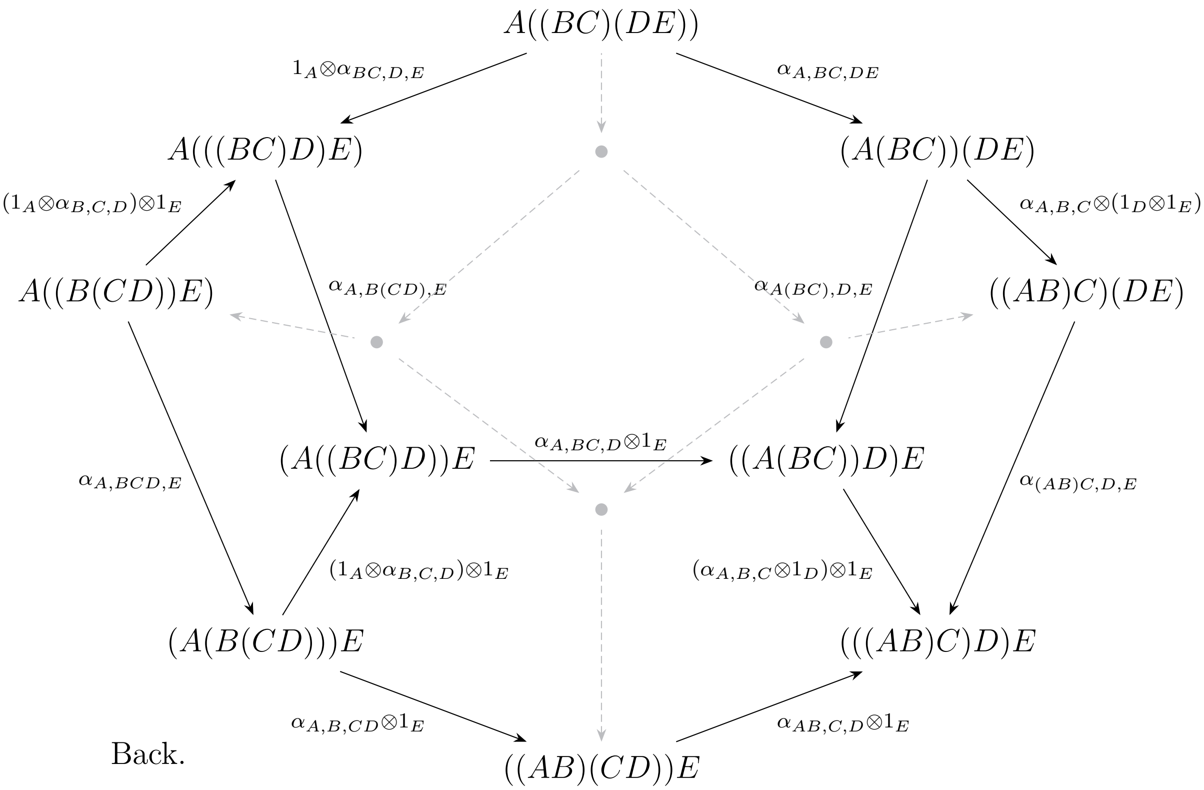
We also know from the axioms of a monoidal category that the 5 products of 4 objects are related via the diagram consisting of natural isomorphisms as below.

$$\begin{tikzcd}
A \otimes (B \otimes (C \otimes D)) \arrow[r, "\alpha_{A,B,C\otimes D}"] \arrow[d, "1_A \otimes \alpha_{B,C,D}"'] & (A \otimes B) \otimes (C \otimes D) \arrow[r, "\alpha_{A\otimes B,C,D}"] & ((A \otimes B) \otimes C) \otimes D \\
A \otimes ((B \otimes C) \otimes D) \arrow[rr, "\alpha_{A,B\otimes C,D}"'] & & (A \otimes (B \otimes C)) \otimes D \arrow[u, "\alpha_{A,B,C} \otimes 1_D"']
\end{tikzcd}$$

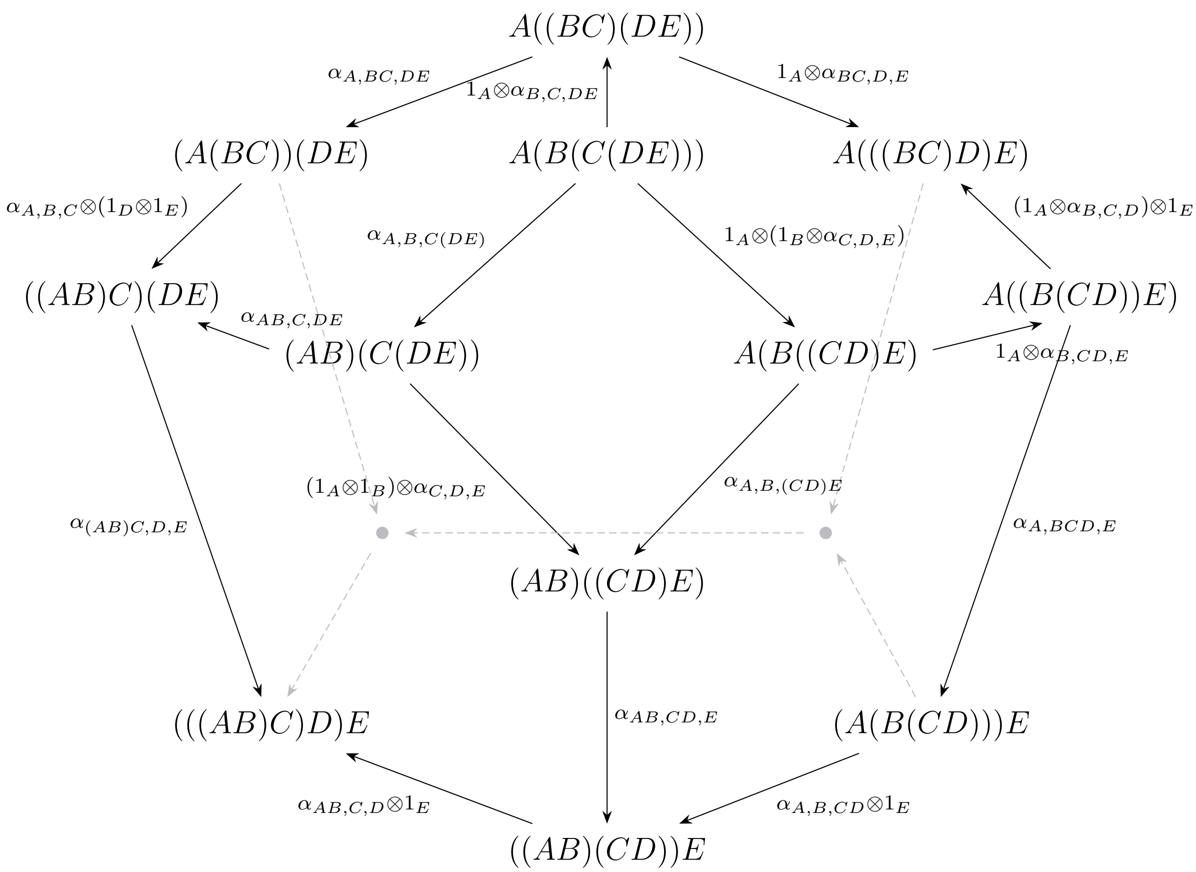
Moreover, this diagram is guaranteed to be commutative for all $A$,$B$,$C$,$D$ in $\mathcal{M}$ (we will elaborate why this is a profound, useful fact).

Finally, repeatedly using instances of $\alpha$, the 14 ways to multiply 5 objects are related via the 3 dimensional diagram as below.

Front. (Note that the symbol $\otimes$ has been suppressed.)




Back.



However, it is not an axiom of monoidal categories that this last diagram is commutative (with a ton of work, one could prove it to be commutative).

To understand what's going on, let us first understand why commutativity is important. The axioms of a monoidal category grant us the commutativity of the pentagon, which connects the five different ways of multiplying four objects $A, B, C, D$. This tells us the following principle: while there are 5 different ways we can multiply four objects $A, B, C, D$, each such choice is **canonically** isomorphic to any other choice.

To see this, suppose you and I want to multiply objects $A, B, C, D$ together. Suppose my favorite way to do it is $(A \otimes B) \otimes (C \otimes D)$, while you choose $(A \otimes (B \otimes C)) \otimes D$. Then we might be in trouble: I have two possible ways, displayed below in blue and orange, to "reparenthesize" my product to get your object.

$$\begin{tikzcd}
A \otimes (B \otimes (C \otimes D)) \arrow[d, "1 \otimes \alpha"'] & (A \otimes B) \otimes (C \otimes D) \arrow[l, "\alpha^{-1}"'] \arrow[r, "\alpha"] & ((A \otimes B) \otimes C) \otimes D \arrow[d, "\alpha^{-1} \otimes 1"] \\
A \otimes ((B \otimes C) \otimes D) \arrow[rr, "\alpha"'] & & (A \otimes (B \otimes C)) \otimes D
\end{tikzcd}$$

Fortunately, the commutativity of the pentagonal diagram enures that the two paths are equal. That is,

$$\alpha \circ ((1 \otimes \alpha) \circ \alpha^{-1}) = (\alpha^{-1} \otimes 1) \circ \alpha.$$

so that, in reality, I actually have *one* unique isomorphism (i.e., a canonical isomorphism) from my object to yours, and you can also canonically get from your object to mine by inverting the unique isomorphism.

However, our choice of two different parenthesizations was arbitrary. The commutativity of the entire diagram therefore tells us that any choice of "parenthesizing" $A \otimes B \otimes C \otimes D$, the product of 4 objects in $\mathcal{M}$, is **canonically** isomorphic to any other possible choice. This brings up a few questions.

- What do we mean by "parenthesizing?"
- What about a product with $n$-many objects $A$ for $n > 4$?

We will rigorously specify what we mean by parenthesizing in a bit. To answer the second question, we state that this result holds for $n > 4$; this is one version of the Coherence Theorem.

# 7.4 Mac Lane's Coherence Theorem

## Step One: Category of Binary Words

To begin the proof of the coherence theorem, we need to first state the theorem itself. This task itself is quite laborious, although it is a worthwhile investment to establish clear terminology and notation, especially in writing the proof itself. Our primary tool will be the abstract concept of a binary word.

**Definition 7.4.1.** Let $x_0, x_1$ be two distinct symbols. A **binary word** $w$ is an element defined recursively as follows.

- $x_0$ and $x_1$ are binary words.
- If $u, v$ are binary words, then $(u) \otimes (v)$ is a binary word.

More precisely, a binary word is any element in the free magma $M = F(\{x_0, x_1\})$ generated by $x_0, x_1$, but we will see that the first definition we offered is more useful and transparent.

---

**Example 7.4.2.** Since $x_0, x_1$ are binary words so is the expression $(x_0) \otimes (x_1)$. Similarly, the expressions

$$(x_0) \otimes ((x_0) \otimes (x_1)) \quad ((x_0) \otimes (x_1)) \otimes x_1$$

are binary words.

---

From the previous example, we see that the notation is a bit clunky. On one hand, our definition, which states that $(u) \otimes (v)$ is a binary word if $u, v$ are, is required so that we can logically manage our parentheses. On the other, it makes notation clunky.

To remedy this, we will often omit parentheses. Given an expression of a binary word, we will always omit the parentheses around individual symbols in the expression. With this rule, we have that:

$$(x_0) \otimes (x_1) = x_0 \otimes x_1$$
$$(x_0) \otimes ((x_0) \otimes (x_1)) = x_0 \otimes (x_0 \otimes x_1)$$
$$((x_0) \otimes (x_1)) \otimes (x_1) = (x_0 \otimes x_1) \otimes x_1$$

That is, we keep the parentheses which group together individual products, and throw away the ones which our smart human brains can don't need.

Next, we move onto an important quantity that we will often perform induction on.

**Definition 7.4.3.** We define the **length of a binary word** $w$, denoted as $\mathcal{L}(w)$, recursively as follows.

- $\mathcal{L}(x_0) = 0$ and $\mathcal{L}(x_1) = 1$

- If $w = u \otimes v$ for two binary words $u, v$, we set $\mathcal{L}(w) = \mathcal{L}(u) + \mathcal{L}(v)$.

---

**Example 7.4.4.** The binary words $(x_1 \otimes x_0) \otimes x_1$, $(x_1 \otimes x_1) \otimes x_0$, $(x_0 \otimes (x_1 \otimes x_1)) \otimes x_0$ all have length 2.

---

More informally, the length of binary word is simply the number of $x_1$ symbols that appear in its expression.

---

**Example 7.4.5.** For any binary word $w$, we have that

$$\mathcal{L}(w \otimes x_0) = \mathcal{L}(x_0 \otimes w) = \mathcal{L}(w).$$

If additionally $u, v$ are binary words, we also have that

$$\begin{aligned}\mathcal{L}(u \otimes (v \otimes w)) &= \mathcal{L}(u) + (\mathcal{L}(v) + \mathcal{L}(w))\\ &= (\mathcal{L}(u) + \mathcal{L}(v)) + \mathcal{L}(w)\\ &= \mathcal{L}((u \otimes v) \otimes w).\end{aligned}$$

We will use the observations made in the previous example later in this section.

---

We now demonstrate that these binary words assemble into a category.

**Definition 7.4.6.** The **category of binary words** is the category $\mathcal{W}$ where

**Objects.** All binary words $w$ of length $n = 0, 1, 2, \dots,$

**Morphisms.** For any two binary words $w$ and $v$, we have that

$$\mathrm{Hom}_{\mathcal{W}}(v, w) = \begin{cases} \{\bullet\} & \text{if } v, w \text{ are the same length} \\ \varnothing & \text{otherwise.} \end{cases}$$

where $\{\bullet\}$ denotes the one point set.

What the above definition tells us is that any two binary words share a morphism if and only if they are of the same length. Moreover, they will only ever share *exactly one* morphism. Since there is always at most one morphism between any two objects in $\mathcal{W}$, we see that $\mathcal{W}$ is a thin category. Moreover, it is monoidal. To prove that it is monoidal, we will need the following small lemma.

**Lemma 7.4.7.** The multiplication of binary words extends to a bifunctor $\otimes : \mathcal{W} \times \mathcal{W} \longrightarrow \mathcal{W}$.

*Proof.* First, we explain how $\otimes : \mathcal{W} \times \mathcal{W} \longrightarrow \mathcal{W}$ operates on objects and morphisms. If $(u, v)$

is an object of $\mathcal{W} \times \mathcal{W}$, we set $\otimes(u, v) = u \otimes v$. Next, consider two morphisms in $\mathcal{W}$.

$$\gamma : u \longrightarrow u' \qquad \beta : v \longrightarrow v'.$$

Note that this implies $\mathcal{L}(u) = \mathcal{L}(u')$ an $\mathcal{L}(v) = \mathcal{L}(v')$, which also imply that

$$\mathcal{L}(u \otimes v) = \mathcal{L}(u) + \mathcal{L}(v) = \mathcal{L}(u') + \mathcal{L}(v') = \mathcal{L}(u' \otimes v').$$

Therefore, we define the image of $(\gamma, \beta)$ under the functor, $\otimes(\gamma, \beta)$, which we more naturally denote as $\gamma \otimes \beta$, to be the unique morphism between $u \otimes v \longrightarrow u' \otimes v'$.

We can picture the action of this functor on objects and morphisms more clearly as below.

$$\boxed{\mathcal{W} \times \mathcal{W} \quad (u_1, v_1) \xrightarrow{(\gamma,\beta)} (u_2, v_2)} \quad \text{maps to} \quad \boxed{\mathcal{W} \quad u \otimes v \xrightarrow{\gamma\otimes\beta} u' \otimes v'}$$

In addition, for any $(u, v)$ in $\mathcal{W} \times \mathcal{W}$, the identity morphism $1_{(u,v)} : (u, v) \longrightarrow (u, v)$ is mapped to the identity $1_{u\otimes v} : u \otimes v \longrightarrow u \otimes v$. Finally, to demonstrate that this respects composition, suppose that $(\gamma, \beta)$ is composable with $(\gamma', \beta')$ as below.

$$\boxed{\mathcal{W} \times \mathcal{W} \quad (u_1, v_1) \xrightarrow{(\gamma,\beta)} (u_2, v_2) \xrightarrow{(\gamma',\beta')} (u_3, v_3)}$$

As both $(\gamma', \beta') \otimes (\gamma, \beta)$ and $(\gamma' \circ \gamma) \otimes (\beta' \circ \beta)$ are parallel morphisms acting as $(u_1, v_1) \longrightarrow (u_3, v_3)$, they must be equal because $\mathcal{W}$ is a thin category (and hence parallel morphisms are equal).

Therefore, we see that $\otimes : \mathcal{W} \times \mathcal{W} \longrightarrow \mathcal{W}$ is a bifunctor. ■

We now show that $\mathcal{W}$ assembles into a monoidal category.

**Proposition 7.4.8.** $(\mathcal{W}, \otimes, x_0)$ is a monoidal category with monoidal product $\otimes : \mathcal{W} \times \mathcal{W} \longrightarrow \mathcal{W}$ and identity object $x_0$.

*Proof.* First, we define our product to be given by the bifunctor $\otimes : \mathcal{W} \times \mathcal{W} \longrightarrow \mathcal{W}$. Second, we define our identity object to be $x_0$. With these two conditions we now need to find unitors, an associator, and check that the necessary diagrams commute.

Now as any two binary words of the same length share a *unique* morphism, all morphisms are isomorphisms. Therefore, by Example 7.4, the isomorphisms

$$\begin{aligned} \alpha_{u,v,w} &: u \otimes (v \otimes w) \xrightarrow{\sim} (u \otimes v) \otimes w \\ \lambda_w &: x_0 \otimes w \xrightarrow{\sim} w \\ \rho_w &: w \otimes x_0 \xrightarrow{\sim} w \end{aligned}$$

are forced to exist. Further, these isomorphisms are natural because all diagrams commute in a thin category. In addition, since $\mathcal{W}$ is a thin category, all diagrams commute, and so, in particular, the required diagrams

$$\begin{tikzcd}
u \otimes (x_0 \otimes v) \arrow[rr, "\alpha_{u,x_0,v}"] \arrow[dr, "1_u \otimes \lambda_v"'] & & (u \otimes x_0) \otimes v \arrow[dl, "\rho_u \otimes 1_v"] \\
& u \otimes v &
\end{tikzcd}$$

$$\begin{tikzcd}
u \otimes (v \otimes (w \otimes z)) \arrow[r, "\alpha_{u,v,w\otimes z}"] \arrow[d, "1_u \otimes \alpha_{v,w,z}"'] & (u \otimes v) \otimes (w \otimes z) \arrow[r, "\alpha_{u\otimes v,w,z}"] & ((u \otimes v) \otimes w) \otimes z \\
u \otimes ((v \otimes w) \otimes z) \arrow[rr, "\alpha_{u,v\otimes w,z}"'] & & (u \otimes (v \otimes w)) \otimes z \arrow[u, "\alpha_{u,v,w} \otimes 1_z"']
\end{tikzcd}$$

also commute, so that $(\mathcal{W}, \otimes, x_0)$ satisfies the axioms of a monoidal category. ■

We now make a few important comments on how to interpret $\alpha, \rho$, and $\lambda$.

- Each $\alpha_{u,v,w} : u \otimes (v \otimes w) \longrightarrow (u \otimes v) \otimes w$ can be thought of as an operator which **shifts** the parentheses to the left. Dually, $\alpha^{-1}_{u,v,w}$ shift them to the right.
- Each $\lambda_w : x_0 \otimes w \xrightarrow{\sim} w$ can be thought of as an operator that **removes** an identity from the left. Dually, $\lambda^{-1}_w$ adds an identity to the left.
- Each $\rho_w : w \otimes x_0 \xrightarrow{\sim} w$ can be thought of as an operator that **removes** an identity from the right. Dually, $\rho^{-1}_w$ adds an identity to the right.

Hence, this very primitive monoidal category $\mathcal{W}$ encodes some basic and useful operators on binary words.

## Step Two: Pure Binary Words

In this section we begin discussing a specific subset of binary words, namely the ones which lack an identity $x_0$. As the theorem is quite complex, this initial restriction allows us to develop intuition and some tools that simplify the proof later.

**Definition 7.4.9.** A **pure binary word** $w$ of length $n$ is a binary word $w$ of length $n$ which has no instance the empty word $x_0$.

---

**Example 7.4.10.** The only pure binary word of length 1 is $x_1$. There is also only one pure binary word of length 2, which is $x_1 \otimes x_1$. The pure binary words of length 3 are

$$x_1 \otimes (x_1 \otimes x_1) \qquad (x_1 \otimes x_1) \otimes x_1$$

and the pure binary words of length 4 are as below.

$$x_1 \otimes (x_1 \otimes (x_1 \otimes x_1)) \quad x_1 \otimes ((x_1 \otimes x_1) \otimes x_1) \quad ((x_1 \otimes x_1) \otimes x_1) \otimes x_1$$
$$x_1 \otimes ((x_1 \otimes x_1) \otimes x_1) \quad (x_1 \otimes (x_1 \otimes x_1)) \otimes x_1$$

---

As a side note, we comment that the number of pure binary words of length $n+1$ is the $n$-th Catalan number

$$C_n = \frac{1}{n+1}\binom{2n}{n} \qquad 1,\ 2,\ 5,\ 14,\ 42,\ 132,\ 429,\ \cdots$$

However, we make no critical use of this fact in our proofs. Next, we form a category of pure binary words.

**Definition 7.4.11.** The **category of pure binary words** $\mathcal{W}_{\text{P}}$ is the full subcategory of $\mathcal{W}$ constructed by restricting the objects of $\mathcal{W}$ to its pure binary words.

More explicitly, $\mathcal{W}_{\text{P}}$ is the category defined as:

**Objects.** All pure binary words $w$ of length $n = 0, 1, 2, \dots,$

**Morphisms.** For any two pure binary words $u, v$ of the same length, we have that $\text{Hom}_{\mathcal{W}_A}(u, v) = \{\bullet\}$, the one point set. No other morphisms are allowed.

We now focus on a particular set of morphisms in $\mathcal{W}_{\text{P}}$. Recall that we may think of each $\alpha_{u,w,v}$ as a "shift map"

$$\alpha_{u,w,v} : u \otimes (v \otimes w) \longrightarrow (u \otimes v) \otimes w$$

which makes a single change in the parenthesis of a binary word. However, $\alpha$ itself does not characterize all possible always in which we make a single change of parentheses within a larger, more complex binary word. An example of this is the morphism

$$1_s \otimes \alpha_{u,v,w} : s \otimes (u \otimes (v \otimes w)) \longrightarrow s \otimes ((u \otimes v) \otimes w)$$

which makes an *internal* change of parentheses. As we will need to focus on these more complicated morphisms, we rigorously define them below.

**Definition 7.4.12** ($\alpha$-arrows)**.** A **forward $\boldsymbol{\alpha}$-arrow** of $\mathcal{W}_{\mathrm{P}}$ is a morphism in $\mathcal{W}_{\mathrm{P}}$ which we recursively define as follows.

- For any triple of pure binary words $w_1, w_2, w_3$ in $\mathcal{W}_{\mathrm{P}}$, the morphism

$$\alpha_{w_1,w_2,w_3} : w_1 \otimes (w_2 \otimes w_3) \longrightarrow (w_1 \otimes w_2) \otimes w_3$$

  is a forward $\alpha$-arrow.

- If $\beta : w \longrightarrow w'$ is a forward $\alpha$-arrow, and $u$ is an arbitrary pure binary word, then the morphisms

$$1_u \otimes \beta : u \otimes w \longrightarrow u \otimes w' \qquad \beta \otimes 1_u : w \otimes u \longrightarrow w' \otimes u$$

  are forward $\alpha$-arrows.

We also define a **backward $\boldsymbol{\alpha}$-arrow** to be the inverse of a forward $\alpha$-arrow.

---

**Example 7.4.13.** Below are a few simple examples of $\alpha$-arrows. The first two are forward, while the third is backward.

$$\begin{tikzcd} x_1 \otimes (x_1 \otimes x_1) \arrow[d, "\alpha_{x_1,x_1,x_1}"'] \\ (x_1 \otimes x_1) \otimes x_1 \end{tikzcd} \quad \begin{tikzcd} x_1 \otimes (x_1 \otimes (x_1 \otimes x_1)) \arrow[d, "1_{x_1} \otimes \alpha_{x_1,x_1,x_1}"'] \\ x_1 \otimes ((x_1 \otimes x_1) \otimes x_1) \end{tikzcd} \quad \begin{tikzcd} (x_1 \otimes x_1) \otimes (x_1 \otimes x_1) \arrow[d, "\alpha^{-1}_{x_1,x_1,x_1 \otimes x_1}"] \\ x_1 \otimes (x_1 \otimes (x_1 \otimes x_1)) \end{tikzcd}$$

We can have even more complicated examples; for example, the morphism below

$$\begin{tikzcd} (u \otimes (x_1 \otimes (x_1 \otimes x_1))) \otimes v \arrow[d, "(1_u \otimes \alpha_{x_1,x_1,x_1}) \otimes 1_v"'] \\ (u \otimes ((x_1 \otimes x_1) \otimes x_1) \otimes v \end{tikzcd}$$

is an $\alpha$-morphism for any pure binary words $u, v$. For example, setting $u = (x_1 \otimes x_1) \otimes x_1$ and $v = x_1 \otimes x_1$, we obtain the forward $\alpha$-arrow as below.

$$\begin{tikzcd} ((x_1 \otimes x_1) \otimes x_1 \otimes (x_1 \otimes (x_1 \otimes x_1))) \otimes (x_1 \otimes x_1) \arrow[d, "(1_{(x_1 \otimes x_1) \otimes x_1} \otimes \alpha_{x_1,x_1,x_1}) \otimes 1_{(x_1 \otimes x_1)}"'] \\ ((x_1 \otimes x_1) \otimes x_1 \otimes ((x_1 \otimes x_1) \otimes x_1) \otimes (x_1 \otimes x_1) \end{tikzcd}$$

---

We emphasize that $\alpha$-arrows only ever involve a single instance of $\alpha$ or $\alpha^{-1}$ in their expression.

Next, we introduce a particularly important instance of a pure binary word that will become essential to our proof.

**Definition 7.4.14.** We define the **terminal word** $w^{(n)}$ of length $n$ recursively as follows.

- $x_1$ is the terminal word of length 1.
- If $w^{(k)}$ is the terminal word of length $k$, then $w^{(k+1)} = w^k \otimes x_1$ is the terminal word of length $k+1$.

More informally, the terminal word is the unique pure binary word of length $n$ for which all parentheses begin on the left.

---

**Example 7.4.15.** Below we list the terminal words by length.

| Length | Terminal Word |
|---|---|
| 1 | $x_1$ |
| 2 | $x_1 \otimes x_1$ |
| 3 | $(x_1 \otimes x_1) \otimes x_1$ |
| 4 | $((x_1 \otimes x_1) \otimes x_1) \otimes x_1$ |
| 5 | $(((x_1 \otimes x_1) \otimes x_1) \otimes x_1) \otimes x_1$ |

---

We now introduce a quantity which provides a "distance-measure" between a pure binary word of length $n$ and the terminal word $w^{(n)}$.

**Definition 7.4.16.** We (recursively) define the **rank** of a binary word as follows.

- $r(x_1) = 0$.
- For a pure binary word of the form $w = u \otimes v$, we set

$$r(u \otimes v) = r(u) + r(v) + \mathcal{L}(v) - 1.$$

---

**Example 7.4.17.** We compute the ranks on the pure binary words of length 4.

$$r(x_1(x_1(x_1x_1))) = 3 \qquad r(x_1((x_1x_1)x_1)) = 2$$
$$r((x_1x_1)(x_1x_1)) = 1 \qquad r((x_1(x_1x_1))x_1) = 1$$
$$r(((x_1x_1)x_1)x_1) = 0$$

---

Note that $w^{(4)} = ((x_1x_1)x_1)x_1$ and $r(((x_1x_1)x_1)x) = 0$. Hence we see that our intuition of the rank being a distance measure from $w^{(n)}$ so far makes sense.

An important property of distance-measuring functions is nonnegativity, which we will now see is satisfied by the rank function.

**Lemma 7.4.18.** Let $w$ be a pure binary word of length $n$. Then $r(w) \geq 0$.

*Proof.* We prove this by induction on $n$. First observe that this clearly holds for $n = 0$ since $r(x_1) = 0$.

Now let $w$ be a pure binary word of length $k$, and suppose the statement is true for all pure binary words with length less than $k$. Since $k > 1$, we may write $w = u \otimes v$ for some pure binary words $u, v$, in which case

$$r(w) = \overbrace{r(u) + r(v)}^{\geq 0 \text{ by induction}} + \mathcal{L}(v) - 1.$$

Since $\mathcal{L}(v) \geq 1$, we see that $r(w) \geq 0$ as desired. ∎

Keeping with the analogy of the rank being a distance measure, we ought to verify that it is zero if and only if the input, which is being measured from $w^{(n)}$, is $w^{(n)}$ itself. We verify that this is the case for the rank function.

**Proposition 7.4.19.** Let $w$ be a pure binary word of length $n$. Then $r(w) = 0$ if and only if $w = w^{(n)}$.

*Proof.* We proceed by induction. In the simplest case, when $n = 1$, we have that $r(x_1) = 0$ by definition. As $x_1 = w^{(1)}$, we see that this satisfies the statement.

Let $w$ be a pure binary word of length $k$, and suppose the statement is true for all pure binary words with length less than $k$. Then we may write our word in the form $w = u \otimes v$, and we have that

$$r(w) = r(u) + r(v) + \mathcal{L}(v) - 1.$$

By Lemma 7.4.18 we know that $r(u), r(v) \geq 0$. Therefore, if $\mathcal{L}(v) > 1$ then $r(w) \neq 0$. Hence, consider the case for when $\mathcal{L}(v) = 1$, so that $v = x_1$. Then

$$r(u \otimes v) = r(u) + r(x_1) + \mathcal{L}(x_1) - 1 = r(u)$$

Therefore, $r(w) = 0$ if and only if if $r(u) = 0$. But by induction, this holds if and only if $u = w^{(k-1)}$. So we see that $w = w^{(k-1)} \otimes x_1 = w^{(k)}$, which proves our result for all $n$. ∎

**Lemma 7.4.20.** Let $\beta : v \longrightarrow w$ be a forward $\alpha$-arrow. Then $r(v) < r(w)$. In other words, forward $\alpha$-arrows decrease rank.

*Proof.* To demonstrate this, we perform induction on the structure of forward $\alpha$-arrows.

Our base case is $\beta = \alpha_{u,v,w} : u \otimes (v \otimes w) \xrightarrow{\sim} (u \otimes v) \otimes w$ for some arbitrary words $u, v, w$. With this case, observe that

$$\begin{aligned} r(u \otimes (v \otimes w)) &= r(u) + r(v \otimes w) + \mathcal{L}(v \otimes w) - 1 \\ &= r(u) + (r(v) + r(w) + \mathcal{L}(w) - 1) \\ &\quad + \mathcal{L}(v \otimes w) - 1 \end{aligned}$$

while

$$\begin{aligned} r((u \otimes v) \otimes w) &= r(u \otimes v) + r(w) + \mathcal{L}(w) - 1 \\ &= r(u) + r(v) + r(w) + \mathcal{L}(v) - 1 + r(w) \\ &+ \mathcal{L}(w) - 1. \end{aligned}$$

If we subtract the quantities, we observe that

$$r(u \otimes (v \otimes w)) - r((u \otimes v) \otimes w) = \mathcal{L}(v \otimes w) - \mathcal{L}(w) > 0$$

since $v$ has at least length 1. Therefore $\alpha_{u,v,w}$ decreases length as desired.

Next, we reach our inductive step: let $\beta = 1_u \otimes \gamma : u \otimes v \longrightarrow u \otimes w$ where $\gamma : v \longrightarrow w$ is a forward $\alpha$-arrow for which the statement is already true. In this case we have that

$$r(u \otimes v) = r(u) + r(v) + \mathcal{L}(v) - 1.$$

while

$$r(u \otimes w) = r(u) + r(w) + \mathcal{L}(w) - 1.$$

Since $\mathcal{L}(v) = \mathcal{L}(w)$ and $r(v) > r(w)$, we see that $r(u \otimes v) > r(u \otimes w)$. Therefore, we see that $\beta = 1_u \otimes \gamma$ decreases rank whenever $\gamma$ is a forward $\alpha$-arrow that also decreases rank.

Finally, let $\beta = \gamma \otimes 1_u$ where $\gamma : v \longrightarrow w$ is a forward $\alpha$ arrow for which the statement is already true. Then we may write $\beta : v \otimes u \longrightarrow w \otimes u$ Now observe that

$$r(v \otimes u) = r(v) + r(u) + \mathcal{L}(u) - 1$$

while

$$r(w \otimes u) = r(w) + r(u) + \mathcal{L}(u) - 1.$$

Since $\gamma : v \longrightarrow w$ decreases rank, we see that $r(v) > r(w)$ and therefore $r(v \otimes u) > r(w \otimes u)$, as desired.

This completes the proof by induction, so that the statement is true for all forward $\alpha$-arrows. ∎

Thus what we have on our hands is the following. We know that the rank of word $w$ is zero if and only if $w = w^{(n)}$. Further, we know that applying $\alpha$-arrows to a pure binary word will decrease its rank. In other words, shifting the parentheses of a pure binary word $w$ brings $w$ "closer" to $w^{(n)}$ (whose parentheses are all on the left). Therefore, the rank of a pure binary word gives us a measure for how far a binary word $w$ is away from $w^{(n)}$.

The following lemma demonstrates our interest in the word $w^{(n)}$.

**Proposition 7.4.21.** Let $w$ be a pure binary word of length $n$. If $w \neq w^{(n)}$, then there exists a finite sequence of forward $\alpha$-arrows from $w$ to $w^{(n)}$.

*Proof.* We first show that for every pure binary word $w \neq w^{(n)}$ there exists a forward $\alpha$-arrow $\beta$ with domain $w$. We prove this statement by induction on length.

Observe the result is immediate for $n = 1, 2$. Suppose the result is true for binary words with length less than $n \geq 3$. Let $w$ be a pure binary word with length $n$. Then $w = u \otimes v$, with $u, v$ other pure binary words. We now consider two cases for $u$ and $v$.

(1) The first case is when $\mathcal{L}(v) = 1$, so that $v = x_1$. As $w \neq w^{(n)}$ we know that $u \neq w^{(n-1)}$, and since $u$ has length less than $w$, we see that by induction there exists a forward $\alpha$-arrow $\beta : u \longrightarrow u'$. Using $\beta$, we can construct the forward $\alpha$-arrow

$$\beta \otimes 1_{x_1} : u \otimes x_1 \longrightarrow u' \otimes x_1.$$

Hence $\beta \otimes 1_{x_1}$ is our desired forward $\alpha$-arrow with domain $w$.

(2) The second case is when $\mathcal{L}(v) > 1$. In this case we may write $w = u \otimes (r \otimes s)$. A natural choice for a forward $\alpha$-arrow in this case is simply

$$\alpha_{u,v,s} : u \otimes (r \otimes s) \longrightarrow (u \otimes r) \otimes s$$

so that this case is also satisfied.

As we see, in all cases for $w \neq w^{(n)}$, we can find a forward $\alpha$-arrow with domain $w$. As $\alpha$-arrows decrease rank, and $r(w) = 0$ if and only if $w^{(n)}$, this guarantees a sequence of $\alpha$-arrows from $w$ to $w^{(n)}$, which is what we set out to show. ■

The previous proposition has an immediate, useful corollary. It will be used as one of the building blocks for the next section.

**Corollary 7.4.22.** Every morphism in $\mathcal{W}_{\text{P}}$ can be expressed as a finite composition of $\alpha$-arrows.

*Proof.* Let $v, w$ be arbitrary pure binary words. Denote $\varphi_{v,w} : v \longrightarrow w$ to be the unique morphism from $v$ to $w$. By Proposition 7.4.21 there exists chains of forward $\alpha$-arrows whose composite we denote as $\Gamma_1 : v \longrightarrow w^{(n)}, \Gamma_2 : w \longrightarrow w^{(n)}$. Our situation is pictured below.

$$\begin{tikzcd}
v \arrow[rr, "\varphi_{v,w}"] \arrow[dr, "\Gamma_1"'] & & w \\
& w^{(n)} \arrow[ur, "\Gamma_2^{-1}"'] &
\end{tikzcd}$$

However, $\mathcal{W}_{\text{P}}$ is a thin category, so parallel morphisms must be equal. Therefore

$$\varphi_{v,w} = \Gamma_2^{-1} \circ \Gamma_1.$$

Hence $\varphi_{v,w}$ is a composition of $\alpha$-arrows. As $\varphi_{v,w}$ was arbitrary, we see that every morphism in $\mathcal{W}_{\text{P}}$ is a finite composition of $\alpha$-arrows. ■

What this corollary says is that every morphism in $\mathcal{W}_{\mathrm{P}}$ can be expressed as a composite of forward and backward $\alpha$-arrows. However, we emphasize that there can be many different ways to represent a morphism in $\mathcal{W}_{\mathrm{P}}$ via $\alpha$-arrows. This will be an issue which we discuss later in the next section.

## Step Three: Coherence for $A^{\otimes n}$ in $\alpha$

Using our results from the previous section, we are almost ready to take our first major step in the proof of Mac Lane's Coherence Theorem. Before we do so, we need to introduce terminology to even state the theorem which we will prove in this section. Towards that goal we introduce a few more definitions.

**Definition 7.4.23.** Let $(\mathcal{M}, \alpha, \lambda, \rho, I, \otimes)$ be a monoidal category. For an object $A$ of $\mathcal{M}$, we define the **proxy map** of $A$ to be a *partial* functor

$$(-)_A : \mathcal{W}_{\mathrm{P}} \longrightarrow \mathcal{M}$$

as follows. Note by *partial* functor, we mean a functor defined on all objects of $\mathcal{W}_{\mathrm{P}}$, but only a subset of all morphisms of $\mathcal{W}_{\mathrm{P}}$.

**Objects.** We define the action on objects recursively as follows.

- We set $(x_1)_A = A$.
- For a binary word $w = u \otimes v$, we define

$$(w)_A = (u \otimes v)_A = (u)_A \otimes (v)_A$$

**Morphisms.** We define the partial functor only on $\alpha$-arrows. We do this recursively as follows.

- For $\alpha_{u,v,w}$ with $u, v, w$ as pure binary words, we set:

$$(\alpha_{u,v,w})_A = \alpha_{(u)_A,(v)_A,(w)_A}$$
$$(\alpha^{-1}_{u,v,w})_A = \alpha^{-1}_{(u)_A,(v)_A,(w)_A}$$

- For $1_u \otimes \beta$ and $\beta \otimes 1_u$ with $\beta$ an $\alpha$-arrow, we set:

$$(1_u \otimes \beta)_A = 1_{(u)_A} \otimes (\beta)_A$$
$$(\beta \otimes 1_u)_A = (\beta)_A \otimes 1_{(u)_A}$$

We now introduce the theorem of the section. This theorem is the first major step in the proof of the coherence theorem, and the rest of this section will be dedicated to proving it.

**Theorem 7.4.24** (Coherence in $\alpha$.)**.** Let $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ be a monoidal category. For every object $A$, there exists a unique functor $\Phi_A : \mathcal{W}_P \longrightarrow \mathcal{M}$ which restricts to the proxy map $(-)_A$ on objects and $\alpha$-arrows of $\mathcal{W}_{\mathrm{P}}$.

We address the question the reader most likely has in mind right now: Why did we only define the proxy map on $\alpha$-arrows? Why not define it on all of the morphisms of $\mathcal{W}_{\mathrm{P}}$ to get a functor to begin with? We did this to avoid a potential well-definedness issue, which we now elaborate on.

Let us attempt to naturally extend the proxy map to a functor. With Corollary 7.4.22, it is clear how to proceed on defining $(-)_A$ on general morphisms. Let $\gamma : v \longrightarrow w$ be any morphism in $\mathcal{W}_{\text{P}}$. By Corollary 7.4.22, there exist forward and backward $\alpha$-arrows $\gamma_1, \dots, \gamma_n$ such that

$$\gamma = \gamma_n \circ \cdots \circ \gamma_1.$$

Since the proxy map is in fact defined on $\alpha$-arrows, and since functors preserve composition, we are required to define

$$(\gamma)_A = (\gamma_n)_A \circ \cdots \circ (\gamma_1)_A.$$

However, we need to be careful. Suppose that we can also express $\gamma$ as the finite composition of $\alpha$-morphisms $\delta_1, \dots, \delta_m$.

$$\gamma = \delta_m \circ \cdots \circ \delta_1.$$

While $\gamma_n \circ \cdots \circ \gamma_1 = \delta_m \circ \cdots \circ \delta_1$ because $\mathcal{W}_{\text{P}}$ is a thin category, and therefore parallel morphisms are equal, we have no idea if

$$(\gamma_n)_A \circ \cdots \circ (\gamma_1)_A = (\delta_m)_A \circ \cdots \circ (\delta_1)_A$$

is true in $\mathcal{M}$. That is, we do not know if equivalent morphisms in $\mathcal{W}_{\text{P}}$ are mapped to equal morphisms under the proxy map. Our issue is one of well-definedness.

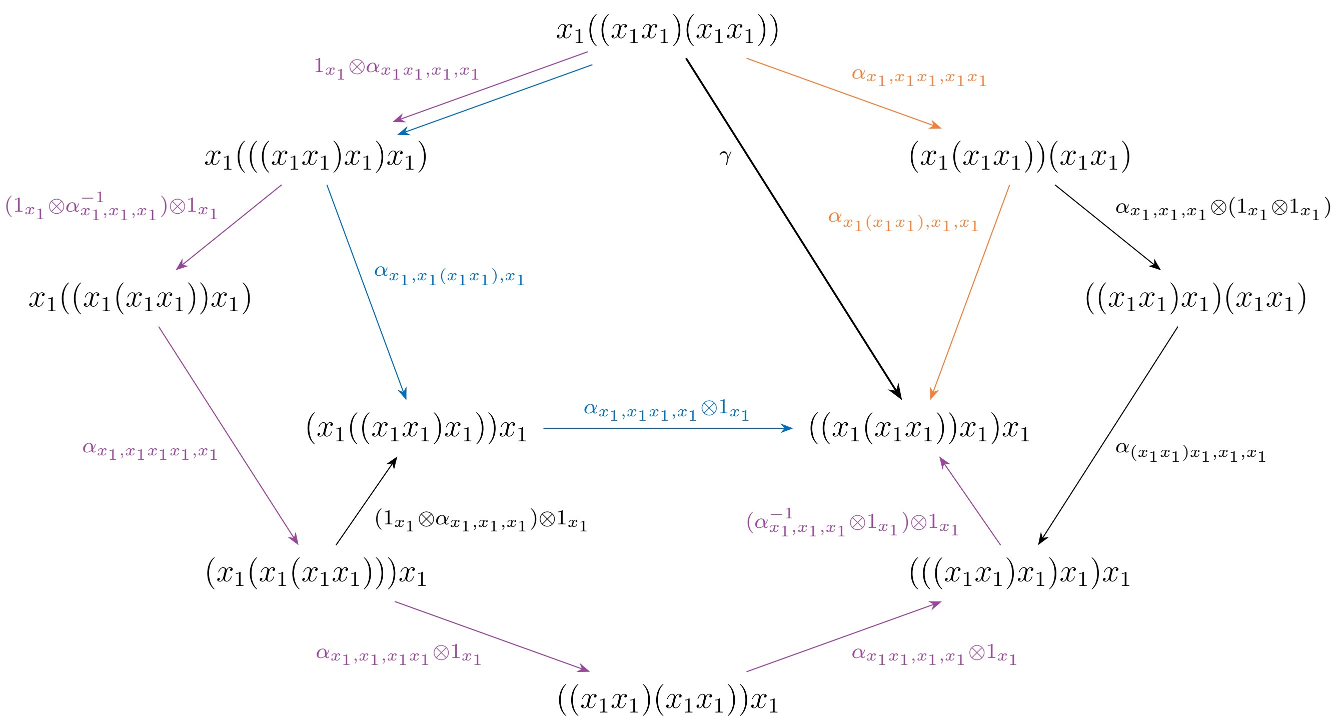
This issue is similar to one which arises in group theory. When one attempts to define a group homomorphism on a quotient group, they must understand that there are different, equivalent ways to represent an element. In this situation they must make sure that the equivalent elements are mapped to the same target in the codomain.

---

**Example 7.4.25.** To illustrate our point, we include a concrete example of our problem which also demonstrates its nontriviality. For notational convenience, we suppress the instances of the monoidal product $\otimes$. Let

$$\gamma : x_1((x_1x_1)(x_1x_1)) \longrightarrow ((x_1(x_1x_1))x_1)x_1.$$

Then we have many possible ways of expressing $\gamma$ in terms of our $\alpha$-arrows. Some potential ways we could express $\gamma$ are displayed below in purple, blue, or orange.




As this is a thin category, we know that the composition of these paths are equal in $\mathcal{W}_{\mathrm{P}}$. However, we now have many ways to define $\gamma$ under the proxy map $(-)_A$. We could write

$$\begin{aligned}(\gamma)_A &= ((\alpha^{-1}_{x_1,x_1,x_1}\otimes 1_{x_1})\otimes 1_{x_1})_A \circ \cdots \circ (1_{x_1}\otimes \alpha_{x_1x_1,x_1,x_1})_A\\ &= (\alpha^{-1}_{A,A,A}\otimes 1_A)\otimes 1_A \circ \cdots \circ 1_A \otimes \alpha_{AA,A,A}\end{aligned}$$

or

$$\begin{aligned}(\gamma)_A &= (\alpha_{x_1,x_1x_1,x_1}\otimes 1_{x_1})_A \circ \cdots \circ (1_{x_1}\otimes \alpha_{x_1x_1,x_1,x_1})_A\\ &= \alpha_{A,AA,A}\otimes 1_A \circ \cdots \circ 1_A \otimes \alpha_{AA,A,A}\end{aligned}$$

or

$$\begin{aligned}(\gamma)_A &= (\alpha_{x_1(x_1x_1),x_1,x_1})_A \circ (\alpha_{x_1,x_1x_1,x_1x_1})_A\\ &= \alpha_{A(AA),A,A}\circ \alpha_{A,AA,AA}\end{aligned}$$

But as morphisms in $\mathcal{M}$, we don't know if these compositions in $\mathcal{M}$, displayed below, are all equal.

$$\begin{tikzcd}
 & & A((AA)(AA)) \arrow[dl, blue, shift right, "1_A\otimes\alpha_{AA,A,A}"'] \arrow[dl, purple, shift left] \arrow[dr, orange, "\alpha_{A,AA,AA}"] \arrow[dddr, "(\gamma)_A"'] & & \\
 & A(((AA)A)A) \arrow[dl, purple, "(1_A\otimes\alpha^{-1}_{A,A,A})\otimes 1_A"'] \arrow[dd, blue, "\alpha_{A,A(AA),A}"] & & (A(AA))(AA) \arrow[dd, orange, "\alpha_{A(AA),A,A}"'] \arrow[dr, "\alpha_{A,A,A}\otimes(1_A\otimes 1_A)"] & \\
A((A(AA))A) \arrow[ddr, purple, "\alpha_{A,AAA,A}"'] & & & & ((AA)A)(AA) \arrow[ddl, "\alpha_{(AA)A,A,A}"] \\
 & (A((AA)A))A \arrow[rr, blue, "\alpha_{A,AA,A}\otimes 1_A"] & & ((A(AA))A)A & \\
 & (A(A(AA)))A \arrow[u, "(1_A\otimes\alpha_{A,A,A})\otimes 1_A"'] \arrow[dr, purple, "\alpha_{A,A,AA}\otimes 1_A"'] & & (((AA)A)A)A \arrow[u, purple, "(\alpha^{-1}_{A,A,A}\otimes 1_A)\otimes 1_A"] & \\
 & & ((AA)(AA))A \arrow[ur, purple, "\alpha_{AA,A,A}\otimes 1_A"'] & &
\end{tikzcd}$$

Hence we need to show that the purple, blue, and orange compositions are equal in $\mathcal{M}$. While we could perform tedious diagram chases to show that they are equal in $\mathcal{M}$, that would only address three of the many possible ways to express $\gamma$. It also would not take care of the case for much larger binary words! Hence, this problem is very nontrivial in general; we need higher level techniques to get what we want.

---

Therefore, to define a functor in the first place, we need to prove the following fact.

**Proposition 7.4.26.** Let $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ be a monoidal category, and let $A$ be an object of $\mathcal{M}$. Let $v, w$ be binary words of the same length. If $\beta_1, \dots, \beta_k$ and $\gamma_1, \dots, \gamma_\ell$ are $\alpha$-arrows with

$$\beta_k \circ \cdots \circ \beta_1, \gamma_\ell \circ \cdots \circ \gamma_1 : v \longrightarrow w$$

then $(\beta_k)_A \circ \cdots \circ (\beta_1)_A = (\gamma_\ell)_A \circ \cdots \circ (\gamma_1)_A$ in $\mathcal{M}$.

To prove this proposition, we will see that it actually suffices to prove the the special case with $w = w^{(n)}$ and with $\beta_1, \dots, \beta_k$ and $\gamma_1, \dots, \gamma_\ell$ all *forward* $\alpha$-arrows. That is, it suffices to prove the following proposition.

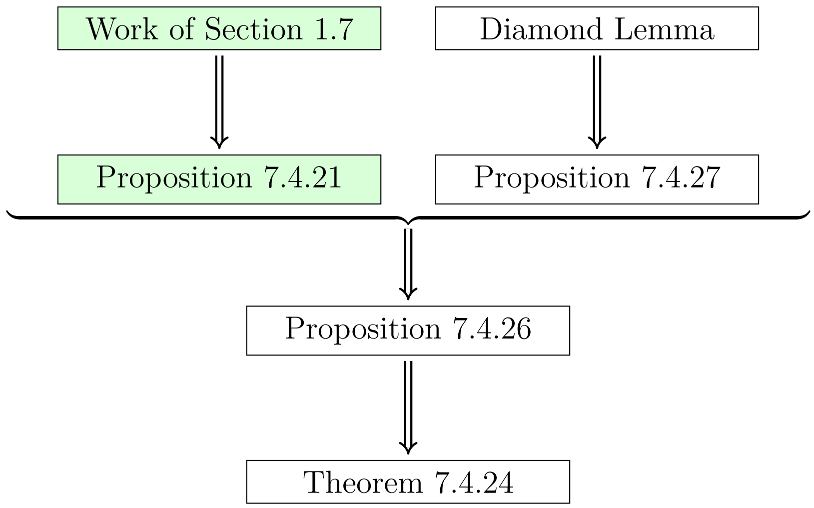
**Proposition 7.4.27.** Let $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ be a monoidal category, and let $A$ be an object of $\mathcal{M}$. Let $w$ be a pure binary word of length $n$. If $\beta_1, \dots, \beta_k$ and $\gamma_1, \dots, \gamma_\ell$ are forward $\alpha$-arrows with

$$\beta_k \circ \cdots \circ \beta_1, \gamma_\ell \circ \cdots \circ \gamma_1 : w \longrightarrow w^{(n)}$$

in $\mathcal{W}_{\mathrm{P}}$, then $(\beta_k)_A \circ \cdots \circ (\beta_1)_A = (\gamma_\ell)_A \circ \cdots \circ (\gamma_1)_A$ in $\mathcal{M}$.

To prove this it will suffice to prove the Diamond Lemma (stated below). It will turn out the bulk of the overall proof toward our theorem will be spent on the Diamond Lemma. At the risk of downplaying its importance, we leave the proof of the Diamond Lemma to the end since it is very tedious and involved, and we do not want to disrupt the flow of the current discussion.

We summarize our plan on how to prove Theorem 7.4.24. The uncolored boxes, and the implications between them, are what is left to do.




**Lemma 7.4.28** (Diamond Lemma)**.** Let $w$ be a pure binary word and suppose $\beta_1, \beta_2$ are two forward $\alpha$-arrows as below.

$$\begin{tikzcd} & w \arrow[dl, "\beta_1"'] \arrow[dr, "\beta_2"] & \\ w_1 & & w_2 \end{tikzcd}$$

There exists a pure binary word $z$ and two $\gamma_1 : w_1 \longrightarrow z, \gamma_2 : w_2 \longrightarrow z$, with $\gamma_1, \gamma_2$ a composition of forward $\alpha$-arrows, such that for any monoidal category $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ the diagram below is commutative in $\mathcal{M}$.

$$\begin{tikzcd} & (w)_A \arrow[dl, "(\beta_1)_A"'] \arrow[dr, "(\beta_2)_A"] & \\ (w_1)_A \arrow[dr, "(\gamma_1)_A"'] & & (w_2)_A \arrow[dl, "(\gamma_2)_A"] \\ & (z)_A & \end{tikzcd}$$

Since the above lemma is an existence result, we emphasize this fact by coloring the arrows, which we are asserting to exist, Green. This is a practice we will continue.

As promised, we now prove Proposition 7.4.27 using the Diamond lemma. We restate the statement of the proposition for the reader's convenience.

**Proposition 7.4.27.** *Let $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ be a monoidal category, and let $A$ be an object of $\mathcal{M}$. Let $w$ be a pure binary word of length $n$. If $\beta_1, \dots, \beta_k$ and $\gamma_1, \dots, \gamma_\ell$ are forward $\alpha$-arrows with*

$$\beta_k \circ \cdots \circ \beta_1, \gamma_\ell \circ \cdots \circ \gamma_1 : w \longrightarrow w^{(n)}$$

*in $\mathcal{W}_{\mathrm{P}}$, then $(\beta_k)_A \circ \cdots \circ (\beta_1)_A = (\gamma_\ell)_A \circ \cdots \circ (\gamma_1)_A$ in $\mathcal{M}$.*

*Proof.* To prove the desired statement, we proceed by induction on the rank of a pure binary word $w$. In what follows we write we will write $w = u \otimes v$ since $\mathcal{L}(w) \geq 3$.

For our base case let $w$ be a word of rank 0. Then by Proposition 7.4.19 we see that $w = w^{(n)}$ so that this statement is trivial.

Next suppose the statement is true for all words with rank at most $k$ where $k \geq 0$. Let $w$ be a pure binary word of rank $k+1$. We want to show that the diagram in $\mathcal{M}$

$$\begin{tikzcd}
& (w)_A \arrow[dl, "(\beta_1)_A"'] \arrow[dr, "(\gamma_1)_A"] & \\
(u_1)_A \arrow[dr, "(\beta_k)_A\circ\cdots\circ(\beta_2)_A"'] & & (v_1)_A \arrow[dl, "(\gamma_\ell)_A\circ\cdots\circ(\gamma_2)_A"] \\
& (w^{(n)})_A &
\end{tikzcd}$$

is commutative. By the Diamond Lemma 7.4.28, there exists exist a pure binary word $z$ and two composites of forward $\alpha$-arrows $\beta'$ and $\gamma'$ such that the diagram below is commutative in $\mathcal{M}$.

$$\begin{tikzcd}
& (w)_A \arrow[dl, "(\beta_1)_A"'] \arrow[dr, "(\gamma_1)_A"] & \\
(u_1)_A \arrow[dr, "(\beta')_A"'] & & (v_1)_A \arrow[dl, "(\gamma')_A"] \\
& (z)_A &
\end{tikzcd}$$

Let $\Gamma_z : z \longrightarrow w^{(n)}$ by any composition of forward $\alpha$-arrows from $z$ to $w^{(n)}$; at least one must exist by Proposition 7.4.21. We can now combine our two diagrams in $\mathcal{M}$ to obtain the diagram below.

$$\begin{tikzcd}
& (w)_A \arrow[dl, "(\beta_1)_A"'] \arrow[dr, "(\gamma_1)_A"] & \\
(u_1)_A \arrow[dr, "(\beta')_A"] \arrow[ddr, "(\beta_k)_A\circ\cdots\circ(\beta_2)_A"'] & & (v_1)_A \arrow[dl, "(\gamma')_A"'] \arrow[ddl, "(\gamma_\ell)\circ\cdots\circ(\gamma_2)_A"] \\
& (z)_A \arrow[d, "(\Gamma_z)_A"] & \\
& w^{(n)} &
\end{tikzcd}$$

By Lemma 7.4.20, we know that forward $\alpha$-arrows decrease rank, so that $r(u_1) < r(w)$ and $r(v_1) > r(w)$. Hence we invoke our induction hypothesis to conclude that both the lower left and lower right triangles commute in $\mathcal{M}$. As the original upper diamond already commutes via the Diamond Lemma, we see that the entire diagram is commutative. Therefore we have that

$$(\beta_k)_A \circ \cdots \circ (\beta_1)_A = (\gamma_\ell)_A \circ \cdots \circ (\gamma_1)_A$$

in $\mathcal{M}$. This completes our induction and hence the proof. ■

As promised, we use the above proposition to prove Proposition 7.4.26.

**Proposition 7.4.26.** *Let $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ be a monoidal category, and let $A$ be an object of $\mathcal{M}$. Let $v, w$ be binary words of the same length. If $\beta_1, \dots, \beta_k$ and $\gamma_1, \dots, \gamma_\ell$ are $\alpha$-arrows with*

$$\beta_k \circ \cdots \circ \beta_1, \gamma_\ell \circ \cdots \circ \gamma_1 : v \longrightarrow w$$

*then $(\beta_k)_A \circ \cdots \circ (\beta_1)_A = (\gamma_\ell)_A \circ \cdots \circ (\gamma_1)_A$ in $\mathcal{M}$.*

*Proof.* We begin by denoting the domain and codomain of the $\alpha$-arrows to make our discussion clear. Let $u_0, \dots, u_k, t_0, \dots, t_\ell$ be the pure binary words such that $u_0 = t_0 = v$, $v_k = u_\ell = w$ and

$$\beta_i : u_{i-1} \longrightarrow u_i, \quad i = 1, 2, \dots, k$$
$$\gamma_j : t_{j-1} \longrightarrow t_j, \quad j = 1, 2, \dots, \ell$$

Note that each morphism may either be forward or backward. With this notation we can picture our parallel $\alpha$-arrows in $\mathcal{W}_{\mathrm{P}}$ as below.

$$\begin{tikzcd}
 & u_1 \arrow[r, "\beta_2"] & u_2 \arrow[r, "\beta_3"] & \cdots \arrow[dr, "\beta_k"] & \\
v \arrow[ur, "\beta_1"] \arrow[dr, "\gamma_1"'] & & & & w \\
 & t_1 \arrow[r, "\gamma_2"'] & t_2 \arrow[r, "\gamma_3"'] & \cdots \arrow[ur, "\gamma_\ell"'] &
\end{tikzcd}$$

Now consider the image of this diagram in $\mathcal{M}$, which we do not yet know to be commutative.

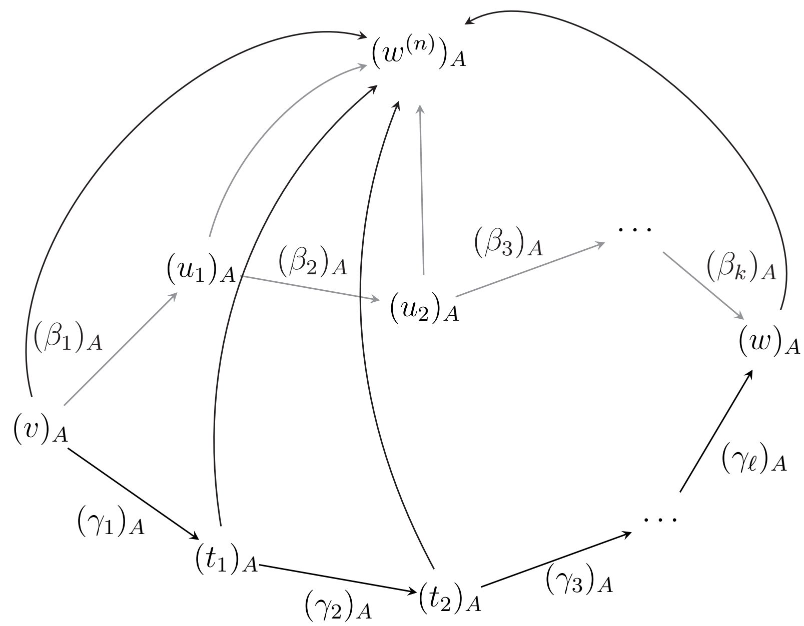
$$
\begin{tikzcd}
 & (u_1)_A \arrow[r, "(\beta_2)_A"] & (u_2)_A \arrow[r, "(\beta_3)_A"] & \cdots \arrow[dr, "(\beta_k)_A"] & \\
(v)_A \arrow[ur, "(\beta_1)_A"] \arrow[dr, "(\gamma_1)_A"'] & & & & (w)_A \\
 & (t_1)_A \arrow[r, "(\gamma_2)_A"'] & (t_2)_A \arrow[r, "(\gamma_3)_A"'] & \cdots \arrow[ur, "(\gamma_\ell)_A"'] &
\end{tikzcd}
$$

Our goal is to show that this diagram in $\mathcal{M}$ is in fact commutative. This will then show our desired equality.

By Proposition 7.4.21, we can connect each pure binary word $u_i$ and $t_i$ to the terminal word $w^{(n)}$ with forward $\alpha$-arrows $\Gamma_{u_i} : u_i \longrightarrow w^{(n)}$ and $\Gamma_{t_i} : t_i \longrightarrow w^{(n)}$. If we add these to our diagram (and suppress the notation on the $\Gamma$'s), it becomes

$$
\begin{tikzcd}
 & & w^{(n)} & & \\
 & u_1 \arrow[r, "\beta_2"] \arrow[ur, bend left] & u_2 \arrow[u] \arrow[r, "\beta_3"] & \cdots \arrow[dr, "\beta_k"] & \\
v \arrow[ur, "\beta_1"] \arrow[dr, "\gamma_1"'] \arrow[uurr, bend left] & & & & w \arrow[uull, bend right] \\
 & t_1 \arrow[r, "\gamma_2"'] \arrow[uuur, bend left] & t_2 \arrow[uuu, bend left] \arrow[r, "\gamma_3"'] & \cdots \arrow[ur, "\gamma_\ell"'] &
\end{tikzcd}
$$

whose image under the proxy map in $\mathcal{M}$ is




Thus the diagram has become a cone, with apex $w^{(n)}$, which is sliced by the triangles. The base of this cone is the original diagram. We now show that each triangle is commutative.

Note that each triangle is of two possible forms: it either consists of $\beta_i$ or $\gamma_i$. Without loss of generality, consider a triangle with an instance of $\beta_i$, as below.

$$\begin{tikzcd}
(u_{i-1})_A \arrow[rr, "(\beta_i)_A"] \arrow[dr, "(\Gamma_{u_{i-1}})_A"'] & & (u_i)_A \arrow[dl, "(\Gamma_{u_i})_A"] \\
& (w^{(n)})_A &
\end{tikzcd}$$

Now if $\beta_i$ is a forward $\alpha$-arrow, observe that by Proposition 7.4.27 it is a commutative diagram in $\mathcal{M}$.

On the other hand, suppose $\beta_i$ is a backward $\alpha$-arrow. Then $\beta_i^{-1}$ is a forward $\alpha$-arrow. Then we may rewrite the triangle as

$$\begin{tikzcd}
(u_{i-1})_A \arrow[dr, "(\Gamma_{u_{i-1}})_A"'] & & (u_i)_A \arrow[ll, "(\beta_i^{-1})_A"'] \arrow[dl, "(\Gamma_{u_i})_A"] \\
& (w^{(n)})_A &
\end{tikzcd}$$

so that it now consists entirely of forward $\alpha$-arrows. This then allows us to apply Proposition 7.4.27 to guarantee that it is a commutative diagram in $\mathcal{M}$. Thus, what we have shown is that each triangle in the above diagram is commutative in $\mathcal{M}$. This literally means that for each $i$,

$$(\Gamma_{u_i})_A \circ (\beta_i)_A = (\Gamma_{u_{i-1}})_A \implies (\beta_i)_A = (\Gamma_{u_i})_A^{-1} \circ (\Gamma_{u_{i-1}})_A$$
$$(\Gamma_{t_i})_A \circ (\gamma_i)_A = (\Gamma_{t_{i-1}})_A \implies (\gamma_i)_A = (\Gamma_{t_i})_A^{-1} \circ (\Gamma_{t_{i-1}})_A$$

Therefore, we see that $(\beta_k)_A \circ \cdots \circ (\beta_1)_A$ can be written as

$$\Big((\Gamma_{u_k})_A^{-1} \circ (\Gamma_{u_{k-1}})_A\Big) \circ \Big((\Gamma_{u_{k-1}})_A^{-1} \circ (\Gamma_{u_{k-2}})_A\Big) \circ \cdots \circ \Big((\Gamma_{u_1})_A^{-1} \circ (\Gamma_{u_0})_A\Big)$$

which is a "telescoping" composition that reduces to

$$(\Gamma_{u_k})_A^{-1} \circ (\Gamma_{u_0})_A.$$

Similarly, we can expression $(\gamma_\ell)_A \circ \cdots \circ (\gamma_1)_A$ as

$$\Big((\Gamma_{t_\ell})_A^{-1} \circ (\Gamma_{t_{\ell-1}})_A\Big) \circ \Big((\Gamma_{t_{\ell-1}})_A^{-1} \circ (\Gamma_{t_{\ell-2}})_A\Big) \circ \cdots \circ \Big((\Gamma_{t_1})_A^{-1} \circ (\Gamma_{t_0})_A\Big)$$

which also reduces to

$$(\Gamma_{t_\ell})_A^{-1} \circ (\Gamma_{t_0})_A.$$

However, $u_k = t_\ell$ and $u_0 = t_0$, so that

$$(\Gamma_{u_k})_A^{-1} \circ (\Gamma_{u_0})_A = (\Gamma_{t_\ell})_A^{-1} \circ (\Gamma_{t_0})_A \implies (\beta_k)_A \circ \cdots \circ (\beta_1)_A = (\beta_k)_A \circ \cdots \circ (\beta_1)_A$$

Thus we have that our original diagram in $\mathcal{M}$

$$\begin{tikzcd}
 & (u_1)_A \arrow[r, "(\beta_2)_A"] & (u_2)_A \arrow[r, "(\beta_3)_A"] & \cdots \arrow[dr, "(\beta_n)_A"] & \\
(v)_A \arrow[ur, "(\beta_1)_A"] \arrow[dr, "(\gamma_1)_A"'] & & & & (w)_A \\
 & (t_1)_A \arrow[r, "(\gamma_2)_A"'] & (t_2)_A \arrow[r, "(\gamma_3)_A"'] & \cdots \arrow[ur, "(\gamma_m)_A"'] &
\end{tikzcd}$$

is commutative. Therefore we have that parallel sequences of $\alpha$-arrows are equal in $\mathcal{M}$, as desired. ■

Finally, we use all of our previous work to prove Theorem 7.4.24. In this case, the proof is simply the definition of our desired functor. We state the theorem here for the reader's convenience.

**Theorem 7.4.24** (Associator Coherence.). *Let $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ be a monoidal category. For every object $A$, there exists a unique functor $\Phi_A : \mathcal{W}_P \longrightarrow \mathcal{M}$ which agrees with the proxy map $(-)_A$ on the objects and $\alpha$-arrows.*

To define this functor, we will (in this order) define the functor on (1) object, (2) $\alpha$-arrows, (3) general morphisms of $\mathcal{W}_{\mathrm{P}}$, and then finally show that our definition preserves composition.
**Objects.** For a pure binary word $w$, we define $\Phi_A(w) = (w)_A$.
**Morphisms.** (1) If $\beta$ is an $\alpha$-arrow, we define $\Phi_A(\beta) = (\beta)_A$.

(2) Now we define our functor on a general morphism $v \longrightarrow w$ in $\mathcal{W}_{\mathrm{P}}$. For convenience denote this as $\varphi_{v,w} : v \longrightarrow w$.

We know by Corollary 7.4.22 that there exist finitely many forward and backward $\alpha$-arrows $\gamma_1, \dots, \gamma_k$ such that

$$\varphi_{v,w} = \gamma_k \circ \cdots \circ \gamma_1.$$

Therefore, define

$$\Phi_A(\varphi_{v,w}) = \Phi(\gamma_k \circ \cdots \circ \gamma_1) = (\gamma_k)_A \circ \cdots \circ (\gamma_1)_A.$$

By Proposition 7.4.26, we see that this definition is well-defined.

Note that this definition allows the functor to also be well-defined on identities, i.e., in all instances, $\Phi_A(1_u) = 1_{u_A}$.

We now show that this definition of our functor behaves under composition. Let $\varphi_{u,v} : u \longrightarrow v$ and $\varphi_{v,w} : v \longrightarrow w$ be morphisms in $\mathcal{W}_{\mathrm{P}}$. Then there exist sequences of $\alpha$-arrows $\beta_1, \dots, \beta_k$ and $\gamma_1, \dots, \gamma_\ell$ such that

$$\varphi_{u,v} = \beta_k \circ \cdots \circ \beta_1 \qquad \varphi_{v,w} = \gamma_\ell \circ \cdots \circ \gamma_1.$$

Then we can write

$$\begin{aligned}
\Phi(\varphi_{v,w} \circ \varphi_{u,v}) &= \Phi(\gamma_\ell \circ \cdots \circ \gamma_1 \circ \beta_k \circ \cdots \circ \beta_1)\\
&= (\gamma_\ell)_A \circ \cdots \circ (\gamma_1)_A \circ (\beta_k)_A \circ \cdots \circ (\beta_1)_A\\
&= \Phi(\gamma_\ell \circ \cdots \circ \gamma_1) \circ \Phi(\beta_k \circ \cdots \circ \beta_1)\\
&= \Phi(\varphi_{v,w}) \circ \Phi(\varphi_{u,v})
\end{aligned}$$

Hence we see that our definition on morphisms behaves appropriately on composition, so that $\Phi$ is in fact a functor.

We conclude this section by proving the Diamond Lemma, which we have now seen to play a critical role in this proof.

**Lemma 7.4.28** (Diamond Lemma). *Let $w$ be a pure binary word and suppose $\beta_1, \beta_2$ are two forward $\alpha$-arrows as below.*

$$
\begin{tikzcd}
 & w \arrow[dl, "\beta_1"'] \arrow[dr, "\beta_2"] & \\
w_1 & & w_2
\end{tikzcd}
$$

*There exists a pure binary word $z$ and two $\gamma_1 : w_1 \longrightarrow z, \gamma_2 : w_2 \longrightarrow z$, with $\gamma_1, \gamma_2$ a composition of forward $\alpha$-arrows, such that for any monoidal category $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ the diagram below is commutative in $\mathcal{M}$.*

$$
\begin{tikzcd}
 & (w)_A \arrow[dl, "(\beta_1)_A"'] \arrow[dr, "(\beta_2)_A"] & \\
(w_1)_A \arrow[dr, "(\gamma_1)_A"'] & & (w_2)_A \arrow[dl, "(\gamma_2)_A"] \\
 & (z)_A &
\end{tikzcd}
$$

*is commutative.*

As we said before, the above lemma is an existence result, so we emphasize this fact by coloring the arrows, which we are asserting to exist, Green.

*Proof.* We will prove this using induction on the length of $w = u \otimes v$. Therefore, throughout the proof, suppose the result is already true for all words of length less than that of $w$.

We proceed in a case-by-case basis, exhausting the possible forms of $\beta_1$ and $\beta_2$. For our purposes, we will express $w = u \otimes v$. Whenever $\mathcal{L}(v) > 1$, we write $v = s \otimes t$.

Let $\beta_1, \beta_2$ be forward $\alpha$-arrows. Then $\beta_1$ could be of the forms

$$\alpha_{u,s,t} \qquad 1_u \otimes \gamma_1 \qquad \gamma_1 \otimes 1_v$$

and $\beta_2$ could be of the forms

$$\alpha_{u,s,t} \qquad 1_u \otimes \gamma_2 \qquad \gamma_2 \otimes 1_v.$$

with $\gamma_1, \gamma_2$ already forward $\alpha$-arrows. Therefore, our cases for $\beta_1, \beta_2$, displayed in tuples, are listed in the table below.

| $(\beta_1, \beta_2)$ | $\alpha_{u,s,t}$ | $1_u \otimes \gamma_2$ | $\gamma_2 \otimes 1_v$ |
|---|---|---|---|
| $\alpha_{u,s,t}$ | $(\alpha_{u,s,t}, \alpha_{u,s,t})$ | $(\alpha_{u,s,t}, 1_u \otimes \gamma_2)$ | $(\alpha_{u,s,t}, \gamma_2 \otimes 1_v)$ |
| $1_u \otimes \gamma_1$ | $(1_u \otimes \gamma_1, \alpha_{u,s,t})$ | $(1_u \otimes \gamma_1, 1_u \otimes \gamma_2)$ | $(1_u \otimes \gamma_1, \gamma_2 \otimes 1_v)$ |
| $\gamma_1 \otimes 1_v$ | $(\gamma_1 \otimes 1_v, \alpha_{u,s,t})$ | $(\gamma_1 \otimes 1_v, 1_u \otimes \gamma_2)$ | $(\gamma_1 \otimes 1_v, \gamma_2 \otimes 1_v)$ |

While there are 9 cases displayed above, we have pointed out via color the pairs of cases which are logically equivalent to each other due to the symmetry of our problem. Therefore, we actually have 6 cases to check We now proceed to the proof.

**Case 1:** $(\alpha_{u,s,t}, \alpha_{u,s,t})$.
In this case, we have that $\beta_1 = \beta_2$, for which the statement is trivially true.

**Case 2:** $(\gamma_1 \otimes 1_v, 1_u \otimes \gamma_2)$
Suppose $\beta_1 = \gamma_1 \otimes 1_v$ and $\beta_2 = 1_u \otimes \gamma_2$. Here, $\gamma_1 : u \longrightarrow u'$ and $\gamma_2 : v \longrightarrow v'$ for some pure binary words $u', v'$. Then we get the diagram

$$\begin{tikzcd}
& u \otimes v \arrow[dl, "\gamma_1\otimes 1_v"'] \arrow[dr, "1_u\otimes\gamma_2"] & \\
u' \otimes v \arrow[dr, "1_{u'}\otimes\gamma_2"'] & & u \otimes v' \arrow[dl, "\gamma_1\otimes 1_{v'}"] \\
& u' \otimes v' &
\end{tikzcd}$$

which commutes by the bifunctoriality of $\otimes$.

**Case 3:** $(\gamma_1 \otimes 1_v, \gamma_2 \otimes 1_v)$
Suppose $\beta_1 = \gamma_1 \otimes 1_v$ and $\beta_2 = \gamma_2 \otimes 1_v$ with $\gamma_1 : u \longrightarrow u_1$ and $\gamma_2 : u \longrightarrow u_2$ both forward $\alpha$-arrows. Then in this case we have the triangle below in $\mathcal{M}$.

$$\begin{tikzcd}
& (u \otimes v)_A \arrow[dl, "(\gamma_1)_A\otimes 1_{(v)_A}"'] \arrow[dr, "(\gamma_2)_A\otimes 1_{(v)_A}"] & \\
(u_1 \otimes v)_A & & (u_2 \otimes v)_A
\end{tikzcd}$$

Note that the above diagram is the image of diagram

$$\begin{tikzcd}
& (u)_A \arrow[dl, "(\gamma_1)_A"'] \arrow[dr, "(\gamma_2)_A"] & \\
(u_1)_A & & (u_2)_A
\end{tikzcd}$$

under the functor $(-) \otimes (v)_A$. As $\mathcal{L}(u) < \mathcal{L}(u \otimes v)$, we know by our induction hypothesis that there exists a pure binary word $z$ and a pair of composite, forward $\alpha$-arrows $\sigma_1 : u_1 \longrightarrow z$ and $\sigma_2 : u_2 \longrightarrow z$ such that the diagram below commutes in $\mathcal{M}$.

$$\begin{tikzcd}
& (u)_A \arrow[dl, "(\gamma_1)_A"'] \arrow[dr, "(\gamma_2)_A"] & \\
(u_1)_A \arrow[dr, "(\sigma_1)_A"'] & & (u_2)_A \arrow[dl, "(\sigma_2)_A"] \\
& (z)_A &
\end{tikzcd}$$

Therefore we can apply the functor $(-)\otimes(v)_A$ on the above diagram to obtain the commutative diagram below

$$
\begin{tikzcd}
 & (u)_A \otimes (v)_A \arrow[dl, "(\beta_2)_A \otimes (1_v)_A"'] \arrow[dr, "(\gamma_1)_A \otimes (1_v)_A"] & \\
(u_1)_A \otimes (v)_A \arrow[dr, "(\sigma_1)_A \otimes (1_v)_A"'] & & (u_2)_A \otimes (v)_A \arrow[dl, "(\sigma_2) \otimes (1_v)_A"] \\
 & (z)_A \otimes (v)_A &
\end{tikzcd}
$$

which proves this case.

**Case 4:** $(1_u \otimes \gamma_1, 1_u \otimes \gamma_2)$
The next case is when $\beta_1 = 1_u \otimes \gamma_1$ and $\beta_2 = 1_u \otimes \gamma_2$ with $\gamma_1 : v \longrightarrow v_1$ and $\gamma_2 : v \longrightarrow v_2$. However, this can be proved in a similar manner as the previous case using the induction hypothesis and the functor $(u)_A \otimes (-)$.
**Case 5:** $(\alpha_{u,s,t}, \gamma_2 \otimes 1_v)$
Let $\beta_1 = \alpha_{u,s,t}$, so that $w = u \otimes (s \otimes t)$. Let $\beta_2 = \gamma_2 \otimes 1_v = \gamma_2 \otimes 1_{s \otimes t}$ with $\gamma_2 : u \longrightarrow u'$ a forward $\alpha$-arrow. Then we will have the diagram in $\mathcal{M}$

$$
\begin{tikzcd}
 & (u \otimes (s \otimes t))_A \arrow[dl, "(\alpha_{u,s,t})_A"'] \arrow[dr, "(\gamma_2 \otimes (1_s \otimes 1_t))_A"] & \\
((u \otimes s) \otimes t)_A \arrow[dr, "((\gamma_2 \otimes 1_s) \otimes 1_t)_A"'] & & (u' \otimes (s \otimes t))_A \arrow[dl, "(\alpha_{u',s,t})_A"] \\
 & ((u' \otimes s) \otimes t)_A &
\end{tikzcd}
$$

which commutes in $\mathcal{M}$ by naturality of $\alpha$.

**Case 6:** $(\alpha_{u,s,t}, 1_u \otimes \gamma_2)$
Let $\beta_1 = \alpha_{u,s,t}$, $\beta_2 = 1_u \otimes \gamma$ with $\gamma$ a forward $\alpha$-arrow with domain $s \otimes t$. By the recursive definition of a forward $\alpha$-arrow, we have three possible cases for $\gamma$.

**Case 6.1:** $\gamma = 1_s \otimes \gamma'$
With $\gamma = 1_s \otimes \gamma'$ with $\gamma' : t \longrightarrow t'$ already a forward $\alpha$-arrow, we have the diagram in $\mathcal{M}$

$$
\begin{tikzcd}
 & (u \otimes (s \otimes t))_A \arrow[dl, "(\alpha_{u,s,t})_A"'] \arrow[dr, "(1_u \otimes (1_s \otimes \gamma))_A"] & \\
((u \otimes s) \otimes t)_A \arrow[dr, "((1_u \otimes 1_s) \otimes \gamma)_A"'] & & (u \otimes (s \otimes t'))_A \arrow[dl, "(\alpha_{u,s,t'})_A"] \\
 & ((u \otimes s) \otimes t')_A &
\end{tikzcd}
$$

which commutes in $\mathcal{M}$ by naturality of $\alpha$.

**Case 6.2:** $\gamma = \gamma' \otimes 1_t$
If $\gamma = \gamma' \otimes 1_t$ with $\gamma' : s \longrightarrow s'$ already a forward $\alpha$-arrow, we can create the diagram

$$\begin{tikzcd}
 & (u \otimes (s \otimes t))_A \arrow[dl, "(\alpha_{u,s,t})_A"'] \arrow[dr, "(1_u \otimes (\gamma' \otimes 1_t))_A"] & \\
((u \otimes s) \otimes t)_A \arrow[dr, "((1_u \otimes \gamma') \otimes 1_t)_A"'] & & (u \otimes (s' \otimes t))_A \arrow[dl, "(\alpha_{u,s',t})_A"] \\
 & ((u \otimes s') \otimes t)_A &
\end{tikzcd}$$

which also commutes in $\mathcal{M}$ by naturality of $\alpha$.
**Case 6.3:** $\gamma = \alpha_{s,p,q}$
The third case for $\gamma$ is when $\gamma = \alpha_{s,p,q}$. In this case, we express $w = u \otimes (s \otimes (p \otimes q))$. We can then construct the diagram

$$\begin{tikzcd}
 & (u \otimes (s \otimes (p \otimes q)))_A \arrow[dl, "(\alpha_{u,s,p\otimes q})_A"'] \arrow[dr, "(1_u \otimes \alpha_{s,p,q})_A"] & \\
((u \otimes s) \otimes (p \otimes q))_A \arrow[d, "(\alpha_{u\otimes s,p,q})_A"'] & & (u \otimes ((s \otimes p) \otimes q))_A \arrow[d, "(\alpha_{u,s\otimes p,q})_A"] \\
(((u \otimes s) \otimes p) \otimes q)_A & & ((u \otimes (s \otimes p)) \otimes q)_A \arrow[ll, "(\alpha_{u,s,p} \otimes 1_q)_A"]
\end{tikzcd}$$

which is always commutative in $\mathcal{M}$. In this case, the word $((u \otimes s) \otimes p) \otimes q$ acts as our vertex $z$ which completes the diagram.

As we have exhausted all possible cases, we see that the statement is true for pure binary words of rank $k+1$ if it is true for all pure binary words with rank at most $k$. By induction, the statement is true for all binary words of any rank, so that we have proved the theorem. ■

## Step Four: Binary Words

So far we have established a unique functor $\Phi_A : \mathcal{W}_{\text{P}} \longrightarrow \mathcal{M}$ for each object $A$ of any given monoidal category $\mathcal{M}$, and this functor grants us coherence in the associators between iterated monoidal products of a single object. We now consider such monoidal products with the identity $I$ as well, so that we may say something about coherence with regard to the unitors $\lambda$ and $\rho$ in a general monoidal category. Towards that goal, we now consider binary words (not just pure binary words) and introduce some definitions.

Recall that $\mathcal{L}$ calculates the length of a binary word, or more informally, the number of $x_1$'s in a binary word. We now introduce a dual quantity which instead counts the number of $x_0$

**Definition 7.4.29.** Let $w$ be a binary word. Define the **identity length** of $w$, denoted $\mathcal{E}$, recursively as follows.

- $\mathcal{E}(x_0) = 1$ and $\mathcal{E}(x_1) = 0$.
- $\mathcal{E}(u \otimes v) = \mathcal{E}(u) + \mathcal{E}(v)$.

Similarly to how $\mathcal{L}(-)$ counts the number of $x_1$'s in a binary word, $\mathcal{E}(-)$ counts the number of $x_0$'s in a binary word.

Next, we introduce the following concept that will later on be key to our proof of Mac Lane's Coherence Theorem.

**Definition 7.4.30.** Let $w$ be a binary word. We define the **clean word** derived from $w$, denoted $\overline{w}$, recursively as follows.

- We set $\overline{x_1} = x_1$.
- If $\mathcal{L}(w) = 0$ (i.e., it has no instance of $x_1$) then $\overline{w} = x_0$.
- Let $u, v$ be binary words with $\mathcal{L}(u) = 0$ and $\mathcal{L}(v) > 0$. Then

$$\overline{u \otimes v} = \overline{v \otimes u} = \overline{v}$$

- Let $u, v$ be binary words with $\mathcal{L}(u), \mathcal{L}(v) > 0$. Then $\overline{u \otimes v} = \overline{u} \otimes \overline{v}$.

Note that for a pure binary word $w$, we have that $\overline{w} = w$. Informally, the clean word of a binary word of nonzero length is simply the pure binary word obtained by removing all instances of the identity from its expression. In the case for a binary word with zero length, we naturally define the clean word to be $x_0$ .

---

**Example 7.4.31.** We offer some examples of clean words obtained from binary words.

| Word | Clean Word |
|---|---|
| $x_0 \otimes (x_0 \otimes x_0)$ | $x_0$ |
| $x_0 \otimes (x_1 \otimes x_0)$ | $x_1$ |
| $(x_1 \otimes x_0) \otimes x_1$ | $x_1 \otimes x_1$ |
| $((x_1 \otimes x_0) \otimes x_0) \otimes x_1$ | $x_1 \otimes x_1$ |
| $(x_1 \otimes x_0) \otimes ((x_1 \otimes x_0) \otimes x_1)$ | $x_1 \otimes (x_1 \otimes x_1)$ |

The above example also shows that two different binary words can have the same clean word.

---

**Definition 7.4.32** (Monoidal Arrows)**.** A **forward monoidal** arrow of $\mathcal{W}$ is defined recursively as follows.

- For any triple of binary words $u, v, w$, the morphisms

$$\begin{aligned} \alpha_{u,v,w} &: u \otimes (v \otimes w) \xrightarrow{\sim} (u \otimes v) \otimes w \\ \lambda_u &: x_0 \otimes u \xrightarrow{\sim} u \\ \rho_u &: u \otimes x_0 \xrightarrow{\sim} u \end{aligned}$$

  are, respectively, **forward** $\alpha$-, $\lambda$-, and $\rho$-**arrows**. They are collectively defined to be **forward monoidal arrows**.

- For any binary word $u$ and forward monoidal arrow $\mu$, the morphisms

$$1_u \otimes \mu \qquad \mu \otimes 1_u$$

  are forward monoidal arrows.

Finally, we say a **backward monoidal arrow** is the inverse of a forward monoidal arrow.

We also establish the following terminology to distinguish our $\alpha$-arrows from our $\lambda$ and $\rho$ arrows.

**Definition 7.4.33.** A **forward unitor arrow** is either a forward $\lambda$-arrow or a forward $\rho$-arrow. Similarly, a **backward unitor arrow** is the inverse of a forward unitor arrow.

As we have already seen forward $\alpha$-arrows, we provide examples of forward and backward $\lambda, \rho$-arrows.

---

**Example 7.4.34.** Below we have a forward and backward $\lambda$-arrow.

$$\begin{tikzcd} x_1 \otimes ((x_0 \otimes x_1) \otimes x_1) \arrow[d, "1_{x_1} \otimes (\lambda_{x_1} \otimes 1_{x_1})"] & (x_1 \otimes x_1) \otimes x_1 \arrow[d, "\lambda^{-1}_{(x_1 \otimes x_1) \otimes x_1}"] \\ x_1 \otimes (x_1 \otimes x_1) & x_0 \otimes ((x_1 \otimes x_1) \otimes x_1) \end{tikzcd}$$

We also have forward and backward $\rho$-arrows below.

$$\begin{tikzcd} (x_1 \otimes x_0) \otimes x_1 \arrow[d, "\rho_{x_1} \otimes 1_{x_1}"] & x_1 \otimes (x_1 \otimes x_1) \arrow[d, "1_{x_1} \otimes \rho^{-1}_{x_1 \otimes x_1}"] \\ x_1 \otimes x_1 & x_1 \otimes ((x_1 \otimes x_1) \otimes x_0) \end{tikzcd}$$

---

We now move onto proving some important lemmas regarding monoidal arrows that we will use for the coherence theorem.

The first three are quick, but have particular importance.

**Lemma 7.4.35.** Let $w$ be a binary word, $w \ne x_0$. Then $\mathcal{E}(w) = 0$ if and only if $w = \overline{w}$.

Note that $w = x_0$ is the only case for which the above proposition is not true, since $x_0 = \overline{x_0}$ but $\mathcal{E}(x_0) \ne 0$. Hence, our reasoning for excluding it (and it is not a case we will need to concern ourselves with).

*Proof.* Suppose $\mathcal{E}(w) = 0$, and let us prove the forward direction by induction on the length of the word. Let us write $w = u \otimes v$, suppose that the statement is true for all pure binary words with length less than $w$. Observe that

$$w = u \otimes v = \overline{u} \otimes \overline{v} = \overline{u \otimes v} = \overline{w}.$$

where we used the induction hypothesis on $u, v$ which have smaller length than $w$. Thus we see that $w = \overline{w}$.

Conversely, suppose $\overline{w} = w$, $w \ne x_0$, and suppose the statement is true for binary words with length less than $w$. Write $w = u \otimes v$. By the definition of a clean word, the only way we can have $\overline{w} = w$ is if $u, v$ are binary words with nonzero length. Therefore, if $\overline{w} = w$ we see that

$$\overline{u} \otimes \overline{v} = u \otimes v.$$

Since $u, v$ have smaller length than $w$, we may use the induction hypothesis to conclude that $\mathcal{E}(u) = \mathcal{E}(v) = 0$. Hence, $\mathcal{E}(w) = 0$, as desired. ∎

**Lemma 7.4.36.** Let $w$ be a binary word. Suppose $\iota : w \longrightarrow w'$ is a forward unitor arrow. Then $\mathcal{E}(w') = \mathcal{E}(w) - 1$.

In other words, any unitor arrow always takes away exactly one identity.

*Proof.* We prove this by examining the possible cases for $\iota$. Write $w = u \otimes v$. As $\iota$ is a forward unitor arrow, it has four possible forms.

(1) Suppose $\iota = \lambda_v : x_0 \otimes v \longrightarrow v$. As

$$\mathcal{E}(v) = \mathcal{E}(v) + \mathcal{E}(x_0) - 1 = \mathcal{E}(v \otimes x_0) - 1$$

we see that the statement is satisfied in this case.

(2) If $\iota = \rho_u : u \otimes x_0 \longrightarrow u$, we can use a similar argument as in (1) to prove the statement.

(3) Suppose $\iota = 1_u \otimes \kappa : u \otimes v \longrightarrow u \otimes v'$ where $\kappa : v \longrightarrow v'$ is a forward unitor arrow for which the statement is already true. Then $\mathcal{E}(v') = \mathcal{E}(v) - 1$. Hence,

$$\mathcal{E}(u \otimes v') = \mathcal{E}(u \otimes v) - 1.$$

Therefore the statement is satisfied for $1_u \otimes \kappa$ if it is true for $\kappa$.

(4) If $\iota = \kappa \otimes 1_v : u \otimes v \longrightarrow u' \otimes v$ where $\kappa$ is a forward unitor for which the statement is already true, then we may prove this case by following a similar argument as in (3).

As we have examined all cases, we may conclude that for every forward unitor $\iota : w \longrightarrow w'$, we have that $\mathcal{E}(w') = \mathcal{E}(w) - 1$ as desired. ■

**Lemma 7.4.37.** Let $\iota : w \longrightarrow w'$ be a forward unitor arrow. Then $\overline{w} = \overline{w}'$.

In other words, unitor arrows do not alter the particular format of a clean word.

*Proof.* First, observe that the result is trivial if $\mathcal{L}(w) = \mathcal{L}(w') = 0$. Therefore, let $w = u \otimes v$ be such a binary word with $\mathcal{E}(w) > 0$. Suppose the statement is true for binary words $v$ such that $\mathcal{E}(v) < \mathcal{E}(w)$. Let $\iota : w \longrightarrow w'$ be a forward unitor arrow. By the recursive definition of $\iota$, our forward unitor arrow has four possible forms.

(1) Suppose $\iota = \lambda_v : x_0 \otimes v \longrightarrow v$. However, note that $\overline{x_0 \otimes v} = \overline{v}$, so that this case is true.

(2) If $\iota = \rho_u : u \otimes x_0 \longrightarrow u$, then this case may be proven in a similar manner as case (1).

(3) Suppose $\iota = 1_u \otimes \kappa : u \otimes v \longrightarrow u \otimes v'$ where $\kappa$ is a forward unitor arrow for which the result is already true. Since $\mathcal{L}(u \otimes v) < 0$, we have a few subcases.

Suppose $\mathcal{L}(v) > 0$. Then by our assumption on $\kappa$, $\overline{v} = \overline{v}'$. Therefore, if $\mathcal{L}(u) = 0$, we see that

$$\overline{u \otimes v} = \overline{v} = \overline{v}' = \overline{u \otimes v}'$$

which satisfies this case. If instead $\mathcal{L}(u) > 0$, then

$$\overline{u \otimes v} = \overline{u} \otimes \overline{v} = \overline{u} \otimes \overline{v}' = \overline{u \otimes v}'$$

which again satisfies the case.

Finally, suppose $\mathcal{L}(v) = 0$. Then $\overline{u \otimes v} = \overline{u} = \overline{u \otimes v}'$.

In all cases we see that $\overline{u \otimes v} = \overline{u \otimes v}'$ as desired.

(3) Our third case if when $\iota = \kappa \otimes 1_v : u \otimes v \longrightarrow u' \otimes v$ with $\kappa$ a forward unitor for which the result is already true. However, this case can be proved similarly as in case (2).

In all instances, we see that for a forward unitor arrow $\iota : w \longrightarrow w'$, we have that $\overline{w} = \overline{w}'$, as desired. ■

The following lemma is an important existence result that will be used in the next proposition.

**Lemma 7.4.38.** Let $w$ be a binary word with $\mathcal{E}(w) > 0$. Then there exists a forward unitor with domain $w$.

*Proof.* We prove this by induction on the total length of a binary word $\mathcal{L}(w) + \mathcal{E}(w)$. Thus, let $w = u \otimes v$ be a binary word with $\mathcal{E}(w) > 0$ and suppose the statement is true for all binary

words $z$ with
$$\mathcal{L}(z) + \mathcal{E}(z) < \mathcal{L}(w) + \mathcal{E}(w).$$

Then we have a few cases for $w$.

(1) Suppose $u = x_0$. Then we take the forward unitor $\lambda_v : x_0 \otimes v \longrightarrow v$.

(2) Suppose $v = x_0$. We may similarly take $\rho_u : u \otimes x_0 \longrightarrow u$, so that this case is satisfied.

(3) Suppose $u, v \neq x_0$. Since $\mathcal{E}(w) > 1$, either $\mathcal{E}(u)$ or $\mathcal{E}(v) > 0$. Without loss of generality, suppose $\mathcal{E}(u) > 0$. Since
$$\mathcal{L}(u) + \mathcal{E}(u) = \mathcal{L}(u) + \mathcal{E}(u)$$
we may apply our induction hypothesis to conclude that there exists a forward unitor $\iota : u \longrightarrow u'$ with domain $u$. Hence, the morphism
$$\iota \otimes 1_v : u \otimes v \longrightarrow u' \otimes v$$
is a forward unitor with domain $u \otimes v = w$.

As we have evaluated all cases, we see that the statement is true for all binary words as desired. ■

The previous four lemmas now give rise to the following proposition.

**Proposition 7.4.39.** Let $w$ be a binary word with $\mathcal{E}(w) = \ell$. Then there exists a composable sequence of $\ell$-many forward unitor arrows $\iota_\ell, \cdots, \iota_1$ as below:
$$\iota_\ell \circ \cdots \circ \iota_1 : w \longrightarrow w'.$$
Moreover, for every such chain, we have that $w' = \overline{w}$.

*Proof.* To prove existence of such a chain for every binary word with nonzero identity length, we may proceed by induction. Let $w$ be a binary word with $\mathcal{E}(w) > 0$, and suppose that such a chain exists for binary words $v$ with $\mathcal{E}(v) < \mathcal{E}(w)$. Then by Lemma 2.5.10, there exists a forward unitor $\iota : w \longrightarrow w'$. By Lemma 2.5.8, $\mathcal{E}(w') = \mathcal{E}(w) - 1$, so by our induction hypothesis, there exists a chain of forward unitor arrows
$$\iota_{\ell-1} \circ \cdots \circ \iota_1 : w' \longrightarrow \overline{w}'.$$
Hence, $\iota \circ \iota_{\ell-1} \circ \cdots \circ \iota_1 : w \longrightarrow \overline{w}$ is a forward chain of unitors with initial domain $w$, which proves existence.

To prove that $w' = \overline{w}$, denote the domain and codomain of our unitors $\iota_i : w_{i-1} \longrightarrow w_i$, so that $w_0 = w$. By Lemma 2.5.9, for each $i$ we have that $\overline{w_{i-1}} = \overline{w_i}$. Hence $\overline{w} = \overline{w_\ell}$. By Lemma 2.5.8, we have that $\mathcal{E}(w_i) = \mathcal{E}(w_{i-1}) - 1$. Therefore,
$$\mathcal{E}(w_\ell) = \mathcal{E}(w) - \ell = 0.$$

However, by Lemma 2.5.7, we see that this implies $w_\ell = \overline{w_\ell} = \overline{w}$. Hence we see that

$$\iota_\ell \circ \cdots \circ \iota : w \longrightarrow \overline{w}$$

as desired. ■

The previous proposition immediately implies the next.

**Proposition 7.4.40.** Let $w$ be a binary word with $\mathcal{L}(w) > 0$. Then there exists a sequence of forward monoidal arrows from $w$ to $w^{(n)}$.

*Proof.* By Lemma 7.4.38, we have a sequence of forward unitor arrows from $w$ to $\overline{w}$.

$$\mu_k \circ \cdots \circ \mu_1 : w \longrightarrow \overline{w}$$

Since $\overline{w}$ is a pure binary word, we can then use Proposition 7.4.21 to guarantee a sequence of forward $\alpha$-arrows from $\overline{w}$ to $w^{(n)}$.

$$\beta_\ell \circ \cdots \circ \beta_1 : \overline{w} \longrightarrow w^{(n)}$$

Composing these morphisms then gives us our desired monoidal arrow:

$$\beta_\ell \circ \cdots \circ \beta_1 \circ \mu_k \circ \cdots \circ \mu_1 : w \longrightarrow w^{(n)}$$

so that such a sequence of forward monoidal arrows exists. ■

And the previous proposition gives us the following corollary.

**Corollary 7.4.41.** Every morphism in $\mathcal{W}$ can be expressed as a composition of a sequence of forward and backward monoidal arrows.

*Proof.* The proof is the same exact proof as that of Corollary 7.4.22. We use the previous proposition with the fact that $\mathcal{W}$ is a thin category to conclude this. ■

## Step Five: Coherence for $A^{\otimes n}$ for $\rho, \lambda$

In this section, we extend the work we've completed with the associators to now include the unitors. We will obtain a theorem similar to Theorem 7.4.24. To even state the theorem, we need to introduce a new definition.

**Definition 7.4.42.** Let $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ be a monoidal category. For each object $A$ in $\mathcal{M}$, we define the **general proxy map** of $A$ to be the partial functor $(-)_A : \mathcal{W} \longrightarrow \mathcal{M}$ defined as follows.

**Objects** We define the general proxy map on objects recursively.

- We set $(x_0)_A = I$ and $(x_1)_A = A$
- For a binary word $w = u \otimes v$ we set:

$$(w)_A = (u \otimes v)_A = (u)_A \otimes (v)_A$$

**Morphisms** We define the partial functor only on $\alpha$-, $\lambda$-, and $\rho$-arrows. This is also done recursively.

- For binary words $u, v, w$, we set:

$$\begin{aligned}
(\alpha_{u,v,w})_A &= \alpha_{(u_A, v_A, w_A)} : u_A \otimes (v_A \otimes w_A) \xrightarrow{\sim} (u_A \otimes v_A) \otimes w_A \\
(\lambda_u)_A &= \lambda_{u_A} : I \otimes u_A \xrightarrow{\sim} u_A \\
(\rho_u)_A &= \rho_{u_A} : u_A \otimes I \xrightarrow{\sim} u_A
\end{aligned}$$

- For a more general $\alpha, \lambda$, or $\rho$-arrow of the form $1_u \otimes \beta$ or $\beta \otimes 1_u$ we set:

$$\begin{aligned}
(1_u \otimes \beta)_A &= 1_{u_A} \otimes (\beta)_A \\
(\beta \otimes 1_u)_A &= (\beta)_A \otimes 1_{u_A}
\end{aligned}$$

Before concluding this definition, we note that there is some potential ambiguity in our definition on the unitors. This is because sometimes a forward unitor arrow in $\mathcal{W}$ can be expressed in two ways. The reader may check that all possible cases for ambiguity are the three cases below.

$$\begin{tikzcd}
x_0 \otimes x_0 \arrow[d, shift right, "\rho_{x_0}"'] \arrow[d, shift left, "\lambda_{x_0}"] \\
x_0
\end{tikzcd}
\qquad
\begin{tikzcd}
x_0 \otimes (x_0 \otimes v) \arrow[d, shift right, "1_{x_0} \otimes \lambda_v"'] \arrow[d, shift left, "\lambda_{(x_0 \otimes v)}"] \\
x_0 \otimes v
\end{tikzcd}
\qquad
\begin{tikzcd}
(u \otimes x_0) \otimes x_0 \arrow[d, shift right, "\rho_u \otimes 1_{x_0}"'] \arrow[d, shift left, "\rho_{(u \otimes x_0)}"] \\
u \otimes x_0
\end{tikzcd}$$

As parallel morphisms in $\mathcal{W}$, they are equal. Therefore, in order for our definition to be well-defined, we need that the corresponding pairs of morphisms

$$
\begin{tikzcd}
I \otimes I \arrow[d, shift right, "\rho_I"'] \arrow[d, shift left, "\lambda_I"] \\
I
\end{tikzcd}
\qquad
\begin{tikzcd}
I \otimes (I \otimes (v)_A) \arrow[d, shift right, "1_I \otimes \lambda_{(v)_A}"'] \arrow[d, shift left, "\lambda_{(I \otimes (v)_A)}"] \\
I \otimes (v)_A
\end{tikzcd}
\qquad
\begin{tikzcd}
((u)_A \otimes I) \otimes I \arrow[d, shift right, "\rho_{(u)_A} \otimes 1_I"'] \arrow[d, shift left, "\rho_{((u)_A \otimes I)}"] \\
(u)_A \otimes I
\end{tikzcd}
$$

to be equal in $\mathcal{M}$. One can show that these morphisms are equal in $\mathcal{M}$ using the unitor diagrams **??**, **??**, and **??**.

Regarding our notation, note that we are recycling the same notation from the proxy map to the general proxy map. This is because the only difference between the two is that the general proxy map is simply an extension of the proxy map which is now defined on identity elements $x_0$ and unitors.

The goal of this section is to prove the following theorem, which can be thought of as an extension of Theorem 7.4.24.

**Theorem 7.4.43** (Coherence in Unitors)**.** Let $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ be a monoidal category. For each object $A$, there exists a unique strict monoidal functor $\Delta_A : \mathcal{W} \longrightarrow \mathcal{M}$ which agrees with the general proxy map on objects and monoidal morphisms.

The above theorem is implied by Proposition 7.4.44 (stated below), in the same way that Theorem 7.4.24 followed from Proposition 7.4.26.

**Proposition 7.4.44.** Let $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ be a monoidal category, and consider two binary words $v, w$. Let $\mu_1, \dots, \mu_k$ and $\eta_1, \dots, \eta_\ell$ be monoidal arrows with:

$$\mu_k \circ \cdots \circ \mu_1, \ \eta_\ell \circ \cdots \circ \eta_1 : v \longrightarrow w$$

Then $(\mu_k)_A \circ \cdots \circ (\mu_1)_A = (\eta_\ell)_A \circ \cdots \circ (\eta_1)_A$ in $\mathcal{M}$.

The above proposition is implied by Proposition 7.4.45 (stated below), in the same way that Proposition 7.4.26 followed from Proposition 7.4.27

**Proposition 7.4.45.** Let $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ be a monoidal category, and consider a binary word $w$. Let $\mu_1, \dots, \mu_k$ and $\eta_1, \dots, \eta_\ell$ be forward monoidal arrows with:

$$\mu_k \circ \cdots \circ \mu_1, \ \eta_\ell \circ \cdots \circ \eta_1 : w \longrightarrow w^{(n)}$$

Then $(\mu_k)_A \circ \cdots \circ (\mu_1)_A = (\eta_\ell)_A \circ \cdots \circ (\eta_1)_A$ in $\mathcal{M}$.

Once we have the above proposition, we can prove Proposition 7.4.44, and hence our desired theorem, using the same technique as in in the Proof of Proposition 7.4.26.

We briefly recall such techniques: We consider two parallel chains of monoidal arrows. We then connect each object in the chain to $w^{(n)}$ with a chain of forward monoidal arrow (recall that a chain must exist for each object). We then have a bunch of adjacent triangles with apex $w^{(n)}$ and we can conclude via the Proposition 7.4.45 that each such triangle commutes. We then conclude that the original two parallel chains form a commutative diagram in $\mathcal{M}$. Thus, our two chains have the same composite in $\mathcal{M}$. This then proves Proposition 7.4.44, which then grants us Theorem 7.4.43.

As our goal has been reduced to proving Proposition 7.4.45, we prove this proposition using the following two results.

The first result is the following proposition.

**Proposition 7.4.46** (Arrow Reorganization)**.** Let $\mu_1, \dots, \mu_k$ be composable forward monoidal arrows with $\ell$-many unitor arrows. Then there exist composable forward unitor arrows $\eta_1, \dots, \eta_\ell$ and forward $\alpha$-arrows $\eta_{\ell+1}, \dots \eta_m$ such that, for any monoidal category $\mathcal{M}$ with object $A$, we have that

$$(\mu_k)_A \circ \cdots \circ (\mu_1)_A = \overbrace{(\eta_m)_A \circ \cdots \circ (\eta_{\ell+1})_A}^{\text{Forward } \alpha' s} \circ \overbrace{(\eta_\ell)_A \circ \cdots \circ (\eta_1)_A}^{\text{Unitors in front}}$$

in $\mathcal{M}$.

The above proposition basically states that monoidal arrows can be reorganized in a particular way with all of the unitors in the front. The second result that we need in order to prove Proposition 7.4.45 is the following proposition.

**Proposition 7.4.47** (Unitor-Chain Equivalence)**.** Let $w$ be a binary word with nonzero length and with $\mathcal{E}(w) = k$. Suppose $\mu_1, \dots, \mu_k$ and $\eta_1, \dots, \eta_k$ are a composable sequence of forward unitor arrows:

$$\mu_k \circ \cdots \circ \mu_1, \ \eta_k \circ \cdots \circ \eta_1 : w \longrightarrow \overline{w}$$

Then $(\mu_k)_A \circ \cdots \circ (\mu_1)_A = (\eta_k)_A \circ \cdots \circ (\eta_1)_A$ in $\mathcal{M}$.

For the sake of organization, we will assume the validity of these two results now so that we may prove 7.4.45 We will then prove these two results in the next section.

***Proof of Proposition 7.4.45***

Let

$$\mu_{n_1} \circ \cdots \circ \mu_1, \ \eta_{n_2} \circ \cdots \circ \eta_1 : w \longrightarrow w^{(n)}$$

be any two composites of forward monoidal arrows from $w$ to $w^{(n)}$. Since $\mathcal{E}(w) = k$ and $\mathcal{E}(w^{(n)}) = 0$, we know by Lemma 7.4.36 that there are exactly $k$-many forward unitors in each expression. We can then use Proposition 7.4.46 to find forward unitor arrows $\gamma_1, \dots \gamma_k, \delta_1, \dots, \delta_k$ and forward $\alpha$-arrows $\gamma_{k+1}, \dots, \gamma_{m_1}, \delta_{k+1}, \dots, \delta_{m_2}$ such that:

$$(\mu_{n_1})_A \circ \cdots \circ (\mu_1)_A = \overbrace{(\gamma_{m_1})_A \circ \cdots \circ (\gamma_{k+1})_A}^{\text{Forward } \alpha' s} \circ \overbrace{(\gamma_k)_A \circ \cdots \circ (\gamma_1)_A}^{\text{Unitors in front}}$$

$$(\eta_{n_2})_A \circ \cdots \circ (\eta_1)_A = \overbrace{(\delta_{m_2})_A \circ \cdots \circ (\delta_{k+1})_A}^{\text{Forward } \alpha' s} \circ \overbrace{(\delta_k)_A \circ \cdots \circ (\delta_1)_A}^{\text{Unitors in front}}$$

By Proposition 7.4.39, we know that the domain of the composition of our unitors is $\overline{w}$:

$$\gamma_k \circ \cdots \circ \gamma_1, \ \delta_k \circ \cdots \circ \delta_1 : w \longrightarrow \overline{w}$$

Diagramatically, our situation is displayed below.

$$\begin{tikzcd}
 & (w)_A \arrow[dl, "(\gamma_1)_A"'] \arrow[dr, "(\delta_1)_A"] & \\
(r_1)_A \arrow[d, "(\gamma_2)_A"'] & & (s_1)_A \arrow[d, "(\delta_2)_A"] \\
\vdots \arrow[dr, "(\gamma_k)_A"'] & & \vdots \arrow[dl, "(\delta_k)_A"] \\
 & (\overline{w})_A \arrow[dl, "(\gamma_{k+1})_A"'] \arrow[dr, "(\delta_{k+1})_A"] & \\
(r_{k+1})_A \arrow[d, "(\gamma_{k+2})_A"'] & & (s_{k+1})_A \arrow[d, "(\delta_{k+2})_A"] \\
\vdots \arrow[dr, "(\gamma_{m_1})_A"'] & & \vdots \arrow[dl, "(\delta_{m_2})_A"] \\
 & \left(w^{(n)}\right)_A &
\end{tikzcd}$$

By Proposition 7.4.50, the upper half of this diagram (above $(\overline{w})_A$) must commute. By Proposition 7.4.26, the bottom half of this diagram (below $(\overline{w})_A$), which consists entirely of forward $\alpha$-arrows, must commute. Therefore, the entire diagram commutes, and this completes the proof. ■

## Step Six: Arrow Reorganization and Unitor Chain Equivalence

We now discuss what it takes to prove the Arrow Reorganization and Unitor-Chain Equivalence results.

To prove the Arrow Reorganization result, it suffices to prove a special case which is precisely stated in the following lemma.

**Lemma 7.4.48** (Associator-Unitor Swap.)**.** Let $\mu : w \longrightarrow w_1$ be a forward $\alpha$-arrow and let $\iota : w_1 \longrightarrow w_2$ be a forward unitor arrow. Then either one of the following two situations must occur.

- There exists a binary word $z$, a forward unitor arrow $\iota' : w \longrightarrow z$ and a forward $\alpha$-arrow $\mu' : z \longrightarrow w_2$ such that, for any monoidal category $\mathcal{M}$, the diagram below commutes.

$$\begin{tikzcd} & (w)_A \arrow[dl, "(\mu)_A"'] \arrow[dr, "(\iota')_A"] & \\ (w_1)_A \arrow[dr, "(\iota)_A"'] & & (z)_A \arrow[dl, "(\mu')_A"] \\ & (w_2)_A & \end{tikzcd}$$

- There exists a forward unitor arrow $\iota' : w \longrightarrow w_2$ such that, for any monoidal category $\mathcal{M}$, the diagram below commutes.

$$\begin{tikzcd} & (w)_A \arrow[dl, "(\mu)_A"'] \arrow[dr, "(\iota')_A"] & \\ (w_1)_A \arrow[rr, "(\iota)_A"'] & & (w_2)_A \end{tikzcd}$$

As before, the above lemma is an existence result, so we emphasize this fact by coloring the arrows that we are asserting to exist Green.

Assuming the above lemma, we prove the Arrow Reorganization Proposition.

***Proof of Arrow Reorganization (Proposition 7.4.46).*** We summarize rather than introducing too much notation, since the proof strategy is rather simple. Consider a sequence of monoidal arrows $\mu_1, \dots, \mu_k$. Suppose $\mu_j$ is a unitor arrow. If $\mu_{j-1}$ is an $\alpha$-arrow, we perform an associator-unitor swap, obtaining a new chain whose composite is the same in $\mathcal{M}$. If not, we leave it alone and check the other unitor arrows.

We perform this reorganization, swapping associator arrows and unitor arrows one at a time, until we have a sequence of morphisms in which no unitor arrow is preceded by an $\alpha$-arrow (and hence all unitors begin at the front of our chain). The repeated application of the Associator-Unitor swap guarantees that the composite of this new chain is equal to the composite of our original chain. ∎

We now understand how to prove the Arrow Reorganization Proposition: it relies critically on the Associator-Unitor Swap. As we now understand how the Associator-Unitor swap is used, we offer its proof.

***Proof of Associator-Unitor Swap (Lemma 7.4.48).*** We prove this using a case-by-case basis. For our proof, we write $w = u \otimes v$. Whenever $\mathcal{L}(v) > 1$, we write $w = u \otimes (s \otimes t)$. If $\mathcal{L}(t) > 1$, we will write $w = u \otimes (s \otimes (p \otimes q))$.

Since $\mu$ is a forward $\alpha$-arrow, it could be of the forms

$$\alpha \quad 1_u \otimes \eta_1 \quad \eta_1 \otimes 1_v$$

with $\eta_1$ a forward $\alpha$-arrow. Since $\iota$ is a forward unitor arrow, it could be of the forms

$$\lambda_v \quad \rho_u \quad 1_u \otimes \eta_2 \quad \eta_2 \otimes 1_v$$

with $\eta_2$ either a forward unitor arrow. We display our table below, this time coloring the entries in order to group together similar cases.

| $(\mu, \iota)$ | $1_u \otimes \eta_2$ | $\eta_2 \otimes 1_v$ | $\lambda_v$ | $\rho_u$ |
|---|---|---|---|---|
| $\alpha$ | $(\alpha_{u,s,t}, 1_u \otimes \eta_2)$ | $(\alpha_{u,s,t}, \eta_2 \otimes 1_v)$ | $(\alpha_{u,s,t}, \lambda_v)$ | $(\alpha_{u,s,t}, \rho_u)$ |
| $1_u \otimes \eta_1$ | $(1_u \otimes \eta_1, 1_u \otimes \eta_2)$ | $(1_u \otimes \eta_1, \eta_2 \otimes 1_v)$ | $(1_u \otimes \eta_1, \lambda_v)$ | $(1_u \otimes \eta_1, \rho_u)$ |
| $\eta_1 \otimes 1_v$ | $(\eta_1 \otimes 1_v, 1_u \otimes \eta_2)$ | $(\eta_1 \otimes 1_v, \eta_2 \otimes 1_v)$ | $(\eta_1 \otimes 1_v, \lambda_v)$ | $(\eta_1 \otimes 1_v, \rho_u)$ |

**Case 1:** $(\alpha_{u,s,t}, 1_{u \otimes s} \otimes \eta_2)$
First consider $\mu = \alpha_{u,s,t} : u \otimes (s \otimes t) \longrightarrow (u \otimes s) \otimes t$ and $\iota = 1_{u \otimes s} \otimes \eta_2$ with $\eta_2 : t \longrightarrow t'$ either a forward $\lambda$ or $\rho$ arrow. We select the forward unitor arrow $1_{u_A} \otimes (1_{s_A} \otimes (\eta_2)_A)$ and the forward $\alpha$-arrow $\alpha_{u_A, s_A, t'_A}$ to obtain the diagram

$$\begin{tikzcd}
 & u_A \otimes (s_A \otimes t_A) \arrow[dl, "\alpha_{u,s,t}"'] \arrow[dr, "1_{u_A} \otimes (1_{s_A} \otimes (\eta_2)_A)"] & \\
(u_A \otimes s_A) \otimes t_A \arrow[dr, "(1_{u_A} \otimes 1_{s_A}) \otimes (\eta_2)_A"'] & & u_A \otimes (s_A \otimes t'_A) \arrow[dl, "\alpha_{u_A, s_A, t'_A}"] \\
 & (u_A \otimes s_A) \otimes t'_A &
\end{tikzcd}$$

which commutes by naturality of $\alpha$.
**Case 2:** $(\alpha_{u,s,t}, \eta_2 \otimes 1_t)$.
In this case, $\mu = \alpha_{u,s,t} : u \otimes (s \otimes t) \longrightarrow (u \otimes s) \otimes t$, while $\iota = \eta_2 \otimes 1_t$. Hence, $\eta_2$ must act on $(u \otimes s)$. With that said, $\eta_2$ must be of the form

$$\lambda_s \qquad \rho_u \qquad \tau \otimes 1_s \qquad 1_u \otimes \sigma$$

with $\tau : u \longrightarrow u'$ and $\sigma : s \longrightarrow s'$ either forward $\lambda$ or $\rho$ arrows. Thus we check each of these cases are satisfied.

**Case 2.1:** $\eta_2 = \lambda_{s_A}$

In this case, $u = I$. We can construct a triangular diagram by appending $\lambda_{s_A \otimes t_A} : I \otimes (s_A \otimes t_A) \longrightarrow s_A \otimes t_A$ as below.

$$\begin{tikzcd} & I \otimes (s_A \otimes t_A) \arrow[dl, "\alpha_{I,s_A,t_A}"'] \arrow[dr, "\lambda_{s_A \otimes t_A}"] & \\ (I \otimes s_A) \otimes t_A \arrow[rr, "\lambda_{s_A} \otimes 1_{t_A}"'] & & s_A \otimes t_A \end{tikzcd}$$

which commutes in $\mathcal{M}$ by Proposition **??**.

**Case 2.2:** $\eta_2 = \rho_u$

In this case, $s_A = I$. We can append the morphism $1_{u_A} \otimes \lambda_{t_A} : u_A \otimes (I \otimes t_A) \longrightarrow u_A \otimes t_A$ to create a triangular diagram as below.

$$\begin{tikzcd} & u_A \otimes (I \otimes t_A) \arrow[dl, "\alpha_{u_A,I,t_A}"'] \arrow[dr, "1_{u_A} \otimes \lambda_{t_A}"] & \\ (u_A \otimes I) \otimes t_A \arrow[rr, "\rho_{u_A} \otimes 1_{t_A}"'] & & u_A \otimes t_A \end{tikzcd}$$

The above diagram is guaranteed to commute by unitor-axiom (Diagram **??**) in any monoidal category $\mathcal{M}$.

**Case 2.3:** $\eta_2 = \tau \otimes 1_s$

In this case, $\eta_2 = \tau \otimes 1_s$ with $\tau$ a forward $\lambda$ or $\rho$-arrow. We can first apply the forward arrow $\tau \otimes (1_{s_A} \otimes 1_{t_A})$ followed by $\alpha_{u'_A,s_A,t_A}$ to obtain the diagram

$$\begin{tikzcd} & u_A \otimes (s_A \otimes t_A) \arrow[dl, "\alpha_{u_A,s_A,t_A}"'] \arrow[dr, "\tau \otimes (1_{s_A} \otimes 1_{t_A})"] & \\ (u_A \otimes s_A) \otimes t_A \arrow[dr, "(\tau \otimes 1_{s_A}) \otimes 1_{t_A}"'] & & u'_A \otimes (s_A \otimes t_A) \arrow[dl, "\alpha_{u'_A,s_A,t_A}"] \\ & (u'_A \otimes s_A) \otimes t_A & \end{tikzcd}$$

which commutes by naturality of $\alpha$.

**Case 2.4:** $\eta_2 = 1_u \otimes \sigma$. This case is nearly identical to the previous, creating a desired diagram which commutes by naturality of $\alpha$.

This proves all of our cases for when $\mu = \alpha_{u_A,s_A,t_A}$ and $\iota = (\eta_2)_A \otimes 1_{t_A}$, and so we move onto our other cases.

**Case 3:** $(\alpha_{u,s,t}, \lambda_t)$

This case cannot happen, since we cannot apply $\lambda : x_0 \otimes t \longrightarrow x_0$ after $\alpha_{u,s,t} : u \otimes (s \otimes t) \longrightarrow (u \otimes s) \otimes t$ as $u \otimes s \neq x_0$ for any binary words $u, s$.

**Case 4:** $(\alpha_{u,s,t}, \rho_{u\otimes s})$
In this case, we'll have that $\mu = \alpha_{u_A,s_A,t_A}$ and $\iota = \rho_{u_A\otimes s_A}$. This implies that $t_A = I$. We can then append the forward $\rho$-arrow $1_{u_A} \otimes \rho_{s_A}$ to obtain the diagram

$$\begin{tikzcd}
 & u_A \otimes (s_A \otimes I) \arrow[dl, "\alpha_{u_A,s_A,I}"'] \arrow[dr, "1_{u_A}\otimes\rho_{s_A}"] & \\
(u_A \otimes s_A) \otimes I \arrow[rr, "\rho_{u_A\otimes s_A}"'] & & u_A \otimes s_A
\end{tikzcd}$$

which we know commutes due to Proposition **??**.
**Case 5:** $(1_u \otimes \eta_1, 1_u \otimes \eta_2)$. In this case $\mu = 1_{u_A} \otimes (\eta_1)_A$ and $\iota = 1_{u_A} \otimes (\eta_2)_A$ with $\eta_1$ a forward $\alpha$-arrow and $\eta_2$ either a forward $\lambda$ or $\rho$-arrow. We can prove this case by induction.

Suppose the statement is true for word of length less than $n$, and let $w = u \otimes v$ be a binary word of length $n$. Then we have the diagram on the left

$$\begin{tikzcd}
 & u_A \otimes v_A \arrow[dl, "1_{u_A}\otimes(\eta_1)_A"'] \\
u_A \otimes v'_A \arrow[dr, "1_u\otimes(\eta_2)_A"'] & \\
 & u_A \otimes v''_A
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
 & v_A \arrow[dl, "(\eta_1)_A"'] \\
v'_A \arrow[dr, "(\eta_2)_A"'] & \\
 & v''_A
\end{tikzcd}$$

which is the image of the diagram on the right under the functor $u_A \otimes (-)$. By induction, there exists either a binary word $z$, and a forward $\lambda$ or $\rho$ arrow $\eta' : v_A \longrightarrow z$ and a forward $\alpha$-arrow $\eta'' : z \longrightarrow v''_A$ such that the diagram below commutes in $\mathcal{M}$.

$$\begin{tikzcd}
 & v_A \arrow[dl, "(\eta_1)_A"'] \arrow[dr, "(\eta')_A"] & \\
v'_A \arrow[dr, "(\eta_2)_A"'] & & z \arrow[dl, "(\eta'')_A"] \\
 & v''_A &
\end{tikzcd}$$

We can then take the image of this under the functor $u_A \otimes (-)$ to obtain the commutative diagram below.

$$\begin{tikzcd}
 & u_A \otimes v_A \arrow[dl, "1_{u_A}\otimes(\eta_1)_A"'] \arrow[dr, "1_{u_A}\otimes(\eta')_A"] & \\
u_A \otimes v'_A \arrow[dr, "1_{u_A}\otimes(\eta_2)_A"'] & & u_A \otimes z \arrow[dl, "1_{u_A}\otimes(\eta'')_A"] \\
 & u_A \otimes v''_A &
\end{tikzcd}$$

As $1_{u_A} \otimes (\eta')_A$ is a forward $\lambda$ or $\rho$ arrow since $(\eta')_A$ is, and since $1_{u_A} \otimes (\eta'')_A$ is a forward $\alpha$-arrow since $(\eta'')_A$ is, we have that the case must be true for all words by induction.

**Case 6:** $(1_u \otimes \eta_1, \eta_2 \otimes 1_{v'})$

In this case, $\mu = 1_{u_A} \otimes (\eta_1)_A$ with $\eta_1 : v \longrightarrow v'$ a forward $\alpha$-arrow, and $\iota = (\eta_2)_A \otimes 1_{v'}$ with $\eta_2 : u \longrightarrow u'$ either a forward $\lambda$ or $\rho$ arrow. We can use the forward $\lambda$ or $\rho$ arrow $(\eta_2)_A \otimes 1_{v_A}$ followed by the $\alpha$-arrow $1_{u'_A} \otimes (\eta_1)_A$ to obtain the diagram below.

$$
\begin{tikzcd}
 & u_A \otimes v_A \arrow[dl, "1_{u_A}\otimes(\eta_1)_A"'] \arrow[dr, "(\eta_2)_A\otimes 1_{v_A}"] & \\
u_A \otimes v'_A \arrow[dr, "(\eta_2)_A\otimes 1_{v'_A}"'] & & u' \otimes v \arrow[dl, "1_{u'_A}\otimes(\eta_1)_A"] \\
 & u'_A \otimes v'_A &
\end{tikzcd}
$$

The above diagram commutes by functoriality of $\otimes$, completing this case.

**Case 7:** $(1_u \otimes \eta_1, \lambda_{v'})$

In this case we'll have $\mu = 1_u \otimes \eta_1$ with $\eta_1$ a forward $\alpha$-arrow and $\iota = \lambda_{v'}$. This then implies that $u = I$. We can then append the $\lambda$-arrow $\lambda_{v_A}$ followed by the $\alpha$-arrow $(\eta_1)_A : v_A \longrightarrow v'_A$ to obtain the diagram

$$
\begin{tikzcd}
 & I \otimes v_A \arrow[dl, "1_I\otimes(\eta_1)_A"'] \arrow[dr, "\lambda_{v_A}"] & \\
I \otimes v'_A \arrow[dr, "\lambda_{v'_A}"'] & & v_A \arrow[dl, "(\eta_1)_A"] \\
 & v'_A &
\end{tikzcd}
$$

which commutes by naturality of $\lambda$.

**Case 8:** $(1_u \otimes \eta_1, \rho_u)$

This case cannot happen, since to apply $\rho_u$ after $1_u \otimes \eta_1$ implies that the codomain of $\eta_1$ is $x_0$, which is not possible if $\eta_1$ is an $\alpha$-morphism.

**Case 9:** $(\eta_1 \otimes 1_v, 1_u \otimes \eta_2)$

Equivalent to **Case 5**.

**Case 10:** $(\eta_1 \otimes 1_v, \eta_2 \otimes 1_v)$

Equivalent to **Case 6**.

**Case 11:**$(\eta_1 \otimes 1_v, \lambda_v)$

This case cannot happen, since to apply $\lambda_v$ after $\eta_1 \otimes 1_v$ implies that the codomain of $\eta_1$ is $x_0$, which is not possible for an $\alpha$-arrow.

**Case 12:** $(\eta_1 \otimes 1_v, \rho_u)$

In this case, we have that $\mu = (\eta_1)_A \otimes 1_{v_A}$ and $\eta_2 = \rho_{u_A}$. This implies that $v_A = I$. We can then append the forward $\rho$ arrow $\rho_{u_A}$ followed by the forward $\alpha$-arrow $(\eta_1)_A$ to the diagram to obtain

$$\begin{tikzcd}
 & u_A \otimes I \arrow[dl, "(\eta_1)_A \otimes 1_{v_A}"'] \arrow[dr, "\rho_{u_A}"] & \\
u'_A \otimes I \arrow[dr, "rho_{u'_A}"'] & & u_A \arrow[dl, "(\eta_1)_A"] \\
 & u'_A &
\end{tikzcd}$$

which commutes by naturality of $\rho$.

This proves all the cases, which completes the proof. ∎

Thus we have proven the Associator-Unitor Swap. Our final task is to prove the Unitor-Chain Equivalence. To do so, it suffices to prove the following lemma.

**Lemma 7.4.49.** (Unitor Diamond Lemma.) Let $w$ be a binary word, and $\mu_1, \mu_2$ a pair of forward unitor arrows as below.

$$\begin{tikzcd}
 & w \arrow[dl, "\mu_1"'] \arrow[dr, "\mu_2"] & \\
w_1 & & w_2
\end{tikzcd}$$

There there exists a binary word $z$ and a pair of forward unitor arrows $\eta_1 : w_1 \longrightarrow z$, $\eta_2 : w_2 \longrightarrow z$ such that for any monoidal category $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$, the diagram below is commutative in $\mathcal{M}$.

$$\begin{tikzcd}
 & (w)_A \arrow[dl, "(\mu_1)_A"'] \arrow[dr, "(\mu_2)_A"] & \\
(w_1)_A \arrow[dr, "(\eta_1)_A"'] & & (w_2)_A \arrow[dl, "(\eta_2)_A"] \\
 & (z)_A &
\end{tikzcd}$$

As before, we color the arrows which we are asserting to exist Green.

*Proof.* To prove this, we do a case-by-case basis again. In general, we will write $w = u \otimes v$, and if $\mathcal{L}(v) > 1$, we write $w = u \otimes (s \otimes t)$.

Now since $\mu_1, \mu_2$ are forward unitor arrows, $\mu_1$ could be of the form

$$1_u \otimes \eta_1 \qquad \eta_1 \otimes 1_v \qquad \lambda_v \qquad \rho_u$$

while $\mu_2$ could be of the form

$$1_u \otimes \eta_2 \qquad \eta_2 \otimes 1_v \qquad \lambda_v \qquad \rho_u$$

with $\eta_1, \eta_2$ both forward unitor arrows. Therefore, our possible cases are as follows. We could have $\mu_1 = \mu_2$. Or, we could have any of the cases below. The paired-coloring indicates logically equivalent cases due to the symmetry of our problem.

| $(\beta_1, \beta_2)$ | $1_u \otimes \eta_2$ | $\eta_2 \otimes 1_v$ | $\lambda_v$ | $\rho_u$ |
|---|---|---|---|---|
| $1_u \otimes \eta_1$ | $(1_u \otimes \eta_1, 1_u \otimes \eta_2)$ | $(1_u \otimes \eta_1, \eta_2 \otimes 1_v)$ | $(1_u \otimes \eta_1, \lambda_v)$ | $(1_u \otimes \eta_1, \rho_u)$ |
| $\eta_1 \otimes 1_v$ | $(\eta_1 \otimes 1_v, 1_u \otimes \eta_2)$ | $(\eta_1 \otimes 1_v, \eta_2 \otimes 1_v)$ | $(\eta_1 \otimes 1_v, \lambda_v)$ | $(\eta_1 \otimes 1_v, \rho_u)$ |
| $\lambda_v$ | $(\lambda_v, 1_u \otimes \eta_2)$ | $(\lambda_v, \eta_2 \otimes 1_v)$ | $(\lambda_v, \lambda_v)$ | $(\lambda_v, \rho_u)$ |
| $\rho_u$ | $(\rho_u, 1_u \otimes \eta_2)$ | $(\rho_u, \eta_2 \otimes 1_v)$ | $(\rho_u, \lambda_v)$ | $(\rho_u, \rho_u)$ |

Since we've already implemented this case-by-case proof strategy several times, we will point out the cases which we've seen before, and take care of the cases that are new.

**Case 1:** $(1_u \otimes \eta_1, 1_u \otimes \eta_2)$ This case can be proven by induction on total length $\mathcal{L}(w) + \mathcal{E}(w)$, using a similar argument as in Case 3 of Lemma 7.4.28.

**Case 2:** $(1_u \otimes \eta_1, \eta_2 \otimes 1_v)$ This case can be proven via functoriality, in a similar manner as Case 2 of Lemma 7.4.28.

**Case 3:** $(1_u \otimes \eta_1, \lambda_v)$.

With $\mu_1 = 1_u \otimes \eta_1$ and $\mu_2 = \lambda_v$, denote $\eta_1 : v \longrightarrow v'$. In this case, we can use the morphisms $\lambda_{(v')_A}$ and $\eta_1$ to obtain the diagram

$$\begin{tikzcd}
 & I \otimes (v)_A \arrow[dl, "1_I \otimes \eta_1"'] \arrow[dr, "\lambda_{(v)_A}"] & \\
I \otimes (v')_A \arrow[dr, "\lambda_{(v')_A}"'] & & v \arrow[dl, "\eta_1"] \\
 & v' &
\end{tikzcd}$$

which commutes by naturality of $\lambda$.

**Case 5:** $(1_u \otimes \eta_1, \rho_u)$.

With $\mu_1 = 1_u \otimes \eta_1, \mu_2 = \rho_u$, note that the only choice for $\eta_1$ is $\eta_1 = 1_{x_0}$. However, there is no unitor arrow with domain $x_0$, so this does not result in a valid case for us to consider.

**Case 6:** $(\eta_1 \otimes 1_v, \lambda_v)$.

With $\mu_1 = \eta_1 \otimes 1_v, \mu_2 = \lambda_v$, note that the only choice for $\eta_1$ is again $1_{x_0}$. Once again, there is no unitor arrow with domain $x_0$, so this is also not a valid case that we need to consider.

**Case 7:** $(\eta_1 \otimes 1_v, \rho_u)$.

With $\mu_1 = \eta_1 \otimes 1_v, \mu_2 = \rho_u$, we can use the morphisms $\rho_{(u')_A}$ and $\eta_1$ to obtain

$$\begin{tikzcd}
 & (u)_A \otimes I \arrow[dl, "\eta_1 \otimes 1_I"'] \arrow[dr, "\rho_{(u)_A}"] & \\
(u')_A \otimes I \arrow[dr, "\rho_{(u')_A}"'] & & (u)_A \arrow[dl, "\eta_1"] \\
 & (u')_A &
\end{tikzcd}$$

which commutes by naturality of $\rho$.

**Case 8:** $(\lambda_v, \lambda_v)$. In this case, we see that $\mu_1 = \mu_2$, so that the statement is trivially satisfied in this case.

With all cases verified, we see that the statement must be true for all binary words, as desired. ■

We now show how this proves the Unitor-Chain Equivalence, which we restate for the readers convenience.

**Proposition 7.4.50** (Unitor-Chain Equivalence)**.** Let $w$ be a binary word with nonzero length and with $\mathcal{E}(w) = k$. Suppose $\mu_1, \dots, \mu_k$ and $\eta_1, \dots, \eta_k$ are forward unitors and that:

$$\mu_k \circ \cdots \circ \mu_1,\ \eta_k \circ \cdots \circ \eta_1 : w \longrightarrow \overline{w}$$

Then $(\mu_k)_A \circ \cdots \circ (\mu_1)_A = (\eta_k)_A \circ \cdots \circ (\eta_1)_A$ in $\mathcal{M}$.

*Proof.* We prove this by induction on $\mathcal{E}(w)$. Suppose the result is true for binary words $v$ with $\mathcal{E}(v) < \mathcal{E}(w)$, and consider two composable chains of forward unitors $\mu_1, \dots, \mu_k, \eta_1, \dots, \eta_k$ as described above. We seek to show that the diagram

$$\begin{tikzcd}
 & (w)_A \arrow[dl, "(\mu_1)_A"'] \arrow[dr, "(\eta_1)_A"] & \\
(u_1)_A \arrow[dr, "(\mu_k)_A\circ\cdots\circ(\mu_2)_A"'] & & (v_1)_A \arrow[dl, "(\eta_k)_A\circ\cdots\circ(\eta_2)_A"] \\
 & (\overline{w})_A &
\end{tikzcd}$$

is commutative in $\mathcal{M}$. By the Unitor Diamond Lemma, there exists a binary word $z$ and two forward unitors $\iota_1 : u \longrightarrow z$ and $\iota_2 : v \longrightarrow z$ such that

$$\begin{tikzcd}
 & (w)_A \arrow[dl, "(\mu_1)_A"'] \arrow[dr, "(\eta_1)_A"] & \\
(u_1)_A \arrow[dr, "(\iota_1)_A"'] & & (v_1)_A \arrow[dl, "(\iota_2)_A"] \\
 & (z)_A &
\end{tikzcd}$$

is commutative in $\mathcal{M}$. Now, by Lemma 7.4.37, we have that $\overline{z} = \overline{w}$. By Lemma 7.4.36, $\mathcal{E}(z) = k - 2$. Hence, by Proposition 7.4.39, there exists a chain of forward unitors $\nu_1, \dots, \nu_{k-2}$ such that $\nu_{k-2} \circ \cdots \circ \nu_1 : z \longrightarrow \overline{w}$. Our situation is displayed below. For clarity, we suppress $\nu_{k-2} \circ \cdots \circ \nu_1 : z \longrightarrow \overline{w}$ in the diagram below.

$$\begin{tikzcd}
 & (w)_A \arrow[dl, "(\mu_1)_A"'] \arrow[dr, "(\eta_1)_A"] & \\
(u_1)_A \arrow[dr, "(\iota_1)_A"] \arrow[ddr, "(\mu_k)_A\circ\cdots\circ(\mu_2)_A"'] & & (v_1)_A \arrow[dl, "(\iota_2)_A"'] \arrow[ddl, "(\eta_k)_A\circ\cdots\circ(\eta_2)_A"] \\
 & (z)_A \arrow[d] & \\
 & (\overline{w})_A &
\end{tikzcd}$$

By Lemma 7.4.36, we know that $\mathcal{E}(u_1), \mathcal{E}(v_1) < \mathcal{E}(w)$. Therefore, we may apply our induction hypothesis to conclude that the lower left and lower right triangles must commute. As the original upper square commutes by the Unitor Diamond Lemma, this implies that

$$(\mu_k)_A \circ \cdots \circ (\mu_1)_A = (\eta_k)_A \circ \cdots \circ (\eta_2)_A$$

as desired. ■

At this point, we have formally filled in all of the potential gaps in the proof of Theorem 7.4.43. We have completed the hard work required to prove Mac Lane's Coherence Theorem. We will use the next section to see how our previous results immediately apply our desired coherence result.

## Step Seven: Proving the Main Theorem

At this point we have proven coherence in associators and unitors, but only when considering iterated monoidal products of a single object. We have not yet achieved our desired result, which should say something about more general monoidal products with different objects in the expression. However, our previous work quickly implies our desired theorem. We first introduce a definition and perform a clever trick.

In what follows, we let $\mathbf{1}$ denote the terminal category whose sole object is denoted $\bullet$.

**Definition 7.4.51.** Let $(\mathcal{M}, \otimes, I)$ be a monoidal category. Define the **iterated functor category[1] of** $\mathcal{M}$, denoted as $\mathbf{It}(\mathcal{M})$, to be the category where:

**Objects.** Functors $F : \mathcal{M}^n \longrightarrow \mathcal{M}$ for all $n = 0, 1, 2, \dots$ When $n = 0$, we let $\mathcal{M}^0 = \mathbf{1}$.

**Morphisms.** Natural transformations $\eta : F \longrightarrow G$ between such functors.

We will give this category a monoidal structure. Towards that goal, we introduce the following bifunctor

$$\odot : \mathbf{It}(\mathcal{M}) \times \mathbf{It}(\mathcal{M}) \longrightarrow \mathbf{It}(\mathcal{M})$$

whose behavior we describe on objects and morphisms as follows.

**On objects.** For two functors $F : \mathcal{M}^n \longrightarrow \mathcal{M}$, $G : \mathcal{M}^m \longrightarrow \mathcal{M}$, we define the functor $F \odot G : \mathcal{M}^{n+m} \longrightarrow \mathcal{M}$ pointwise as

$$(F \odot G)(A_1, \dots, A_{n+m}) = F(A_1, \dots, A_n) \otimes G(A_{n+1}, \dots, A_{n+m})$$

where $\otimes$ is the monoidal product of $\mathcal{M}$.

**On morphisms.** Let $F_1, G_1 : \mathcal{M}^n \longrightarrow \mathcal{M}$ and $F_2, G_2 : \mathcal{M}^m \longrightarrow \mathcal{M}$. Given natural transformations

$$\eta : F_1 \longrightarrow G_1 \qquad \mu : F_2 \longrightarrow G_2$$

we define the natural transformation $\eta \odot \mu : F_1 \odot G_1 \longrightarrow F_2 \odot G_2$ pointwise as

$$(\eta \odot \mu)_{(A_1,\dots,A_{n+m})} = (\eta)_{(A_1,\dots,A_n)} \otimes (\mu)_{(A_{n+1},\dots,A_{n+m})}$$

The above bifunctor is what allows us to regard $\mathbf{It}(\mathcal{M})$ as a monoidal category. This is more precisely stated in the following lemma.

**Lemma 7.4.52.** Let $(\mathcal{M}, \otimes, I, \alpha, \lambda, \rho)$ be a monoidal category. Then

$$(\mathbf{It}(\mathcal{M}), \odot, c, \boldsymbol{\alpha}, \boldsymbol{\lambda}, \boldsymbol{\rho})$$

is a monoidal category where

- The monoidal product is the bifunctor $\odot : \mathbf{It}(\mathcal{M}) \times \mathbf{It}(\mathcal{M}) \longrightarrow \mathbf{It}(\mathcal{M})$
- The identity object is the functor $c : \mathbf{1} \longrightarrow \mathcal{M}$, where $c(\bullet) = I$

[1]The notation of this category is due to Mac Lane, but he did not supply a name for this category. So I made one up. Today, this construction is known as an endomorphism operad.

- For functors $F_j : \mathcal{M}^{i_j} \longrightarrow \mathcal{M}$, $j = 1, 2, 3$, the associator

$$\boldsymbol{\alpha}_{F_1,F_2,F_3} : F_1 \odot (F_2 \odot F_3) \longrightarrow (F_1 \odot F_2) \odot F_3$$

  is the natural transformation defined pointwise for each $(A_1, \dots, A_{i_1+i_2+i_3}) \in \mathcal{M}^{(i_1+i_2+i_3)}$ as

$$(\boldsymbol{\alpha}_{F_1,F_2,F_3})_{(A_1,\dots,A_{i_1+i_2+i_3})} = \alpha_{(F(A_1,\dots,A_{i_1}),F(A_{i_1+1},\dots A_{i_1+i_2}),F(A_{i_1+i_2+1},\dots,A_{i_1+i_2+i_3}))}$$

- For a functor $F : \mathcal{M}^n \longrightarrow \mathcal{M}$, the left unitor $\boldsymbol{\lambda} : c \odot F \longrightarrow F$ is the natural transformation defined pointwise for $(\bullet, A_1, \dots, A_n) \in \mathbf{1} \times \mathcal{M}^n$ as

$$(\boldsymbol{\lambda}_F)_{(\bullet,A_1,\dots,A_n)} = \lambda_{F(A_1,\dots,A_n)}$$

  while the right unitor $\boldsymbol{\rho} : F \odot c \longrightarrow F$ is the natural transformation defined similarly as

$$(\boldsymbol{\rho}_F)_{(A_1,\dots,A_n,\bullet)} = \rho_{F(A_1,\dots,A_n)}$$

It is simple to check that these satisfy the axioms of a monoidal category. We now reach the final theorem.

**Theorem 7.4.53** (Coherence Theorem for Monoidal Categories.)**.** For every monoidal category $\mathcal{M}$, there exists a unique, strict monoidal functor

$$\Phi_{\text{id}} : \mathcal{W} \longrightarrow \mathbf{It}(\mathcal{M})$$

where $\Phi_{\text{id}}(x_1) = \text{id} : \mathcal{M} \longrightarrow \mathcal{M}$.

*Proof.* As $(\mathbf{It}(\mathcal{M}), \odot, c)$ is a monoidal category by Lemma 7.4.52, the theorem follows by a simple application of Theorem 7.4.43 to this monoidal category. ■

A reader might be wondering: How does the above theorem grant us coherence? Let us first investigate the behavior of this functor.

Under the functor, the morphism in $\mathcal{W}$

$$x_1 \otimes (x_1 \otimes x_1) \xrightarrow{\alpha_{x_1,x_1,x_1}} (x_1 \otimes x_1) \otimes x_1$$

is mapped by $\Phi_{\text{id}}$ to the natural transformation between the functors in $\mathbf{It}(\mathcal{M})$

$$\text{id} \odot (\text{id} \odot \text{id}) \xrightarrow{\alpha_{\text{id,id,id}}} (\text{id} \odot \text{id}) \odot \text{id}$$

and, as functors from $\mathcal{M}^3 \longrightarrow \mathcal{M}$, we may substitute any $A, B, C$ to obtain a natural isomorphism

$$\alpha_{A,B,C} : A \otimes (B \otimes C) \longrightarrow (A \otimes B) \otimes C$$

in $\mathcal{M}$. Next, we know that functors preserve diagrams. Therefore, our commutative pentagon diagram in $\mathcal{W}$

$$\begin{tikzcd}
x_1 \otimes (x_1 \otimes (x_1 \otimes x_1)) \arrow[r] \arrow[d] & (x_1 \otimes x_1) \otimes (x_1 \otimes x_1) \arrow[r] & ((x_1 \otimes x_1) \otimes x_1) \otimes x_1 \\
x_1 \otimes ((x_1 \otimes x_1) \otimes x_1) \arrow[rr] & & (x_1 \otimes (x_1 \otimes x_1)) \otimes x_1 \arrow[u]
\end{tikzcd}$$

is mapped by $\Phi_{\mathrm{id}}$ to a commutative diagram of natural transformations in $\mathbf{It}(\mathcal{M})$ between the functors below

$$\begin{tikzcd}
\mathrm{id} \odot (\mathrm{id} \odot (\mathrm{id} \odot \mathrm{id})) \arrow[r] \arrow[d] & (\mathrm{id} \odot \mathrm{id}) \odot (\mathrm{id} \odot \mathrm{id}) \arrow[r] & ((\mathrm{id} \odot \mathrm{id}) \odot \mathrm{id}) \odot \mathrm{id} \\
\mathrm{id} \odot ((\mathrm{id} \odot \mathrm{id}) \odot \mathrm{id}) \arrow[rr] & & (\mathrm{id} \odot (\mathrm{id} \odot) \, \mathrm{id}) \odot \mathrm{id} \arrow[u]
\end{tikzcd}$$

and as the above functors are of the form $\mathcal{M}^4 \longrightarrow \mathcal{M}$, we may substitute any $A, B, C, D \in \mathcal{M}$ to obtain the commutative diagram

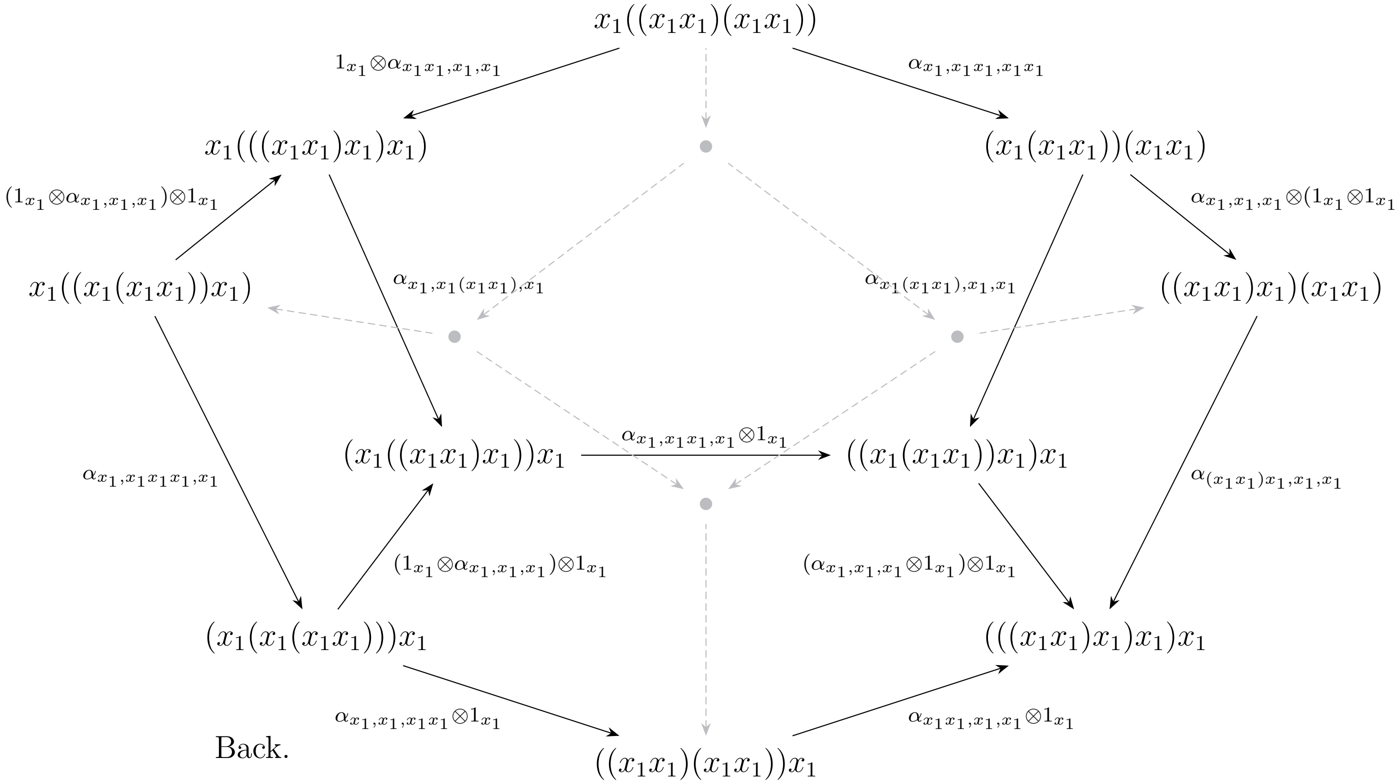
$$\begin{tikzcd}
A \otimes (B \otimes (C \otimes D)) \arrow[r, "\alpha_{A,B,C\otimes D}"] \arrow[d, "1_A \otimes \alpha_{B,C,D}"'] & (A \otimes B) \otimes (C \otimes D) \arrow[r, "\alpha_{A\otimes B,C,D}"] & ((A \otimes B) \otimes C) \otimes D \\
A \otimes ((B \otimes C) \otimes D) \arrow[rr, "\alpha_{A,B\otimes C,D}"'] & & (A \otimes (B \otimes C)) \otimes D \arrow[u, "\alpha_{A,B,C} \otimes 1_D"']
\end{tikzcd}$$

in $\mathcal{M}$.

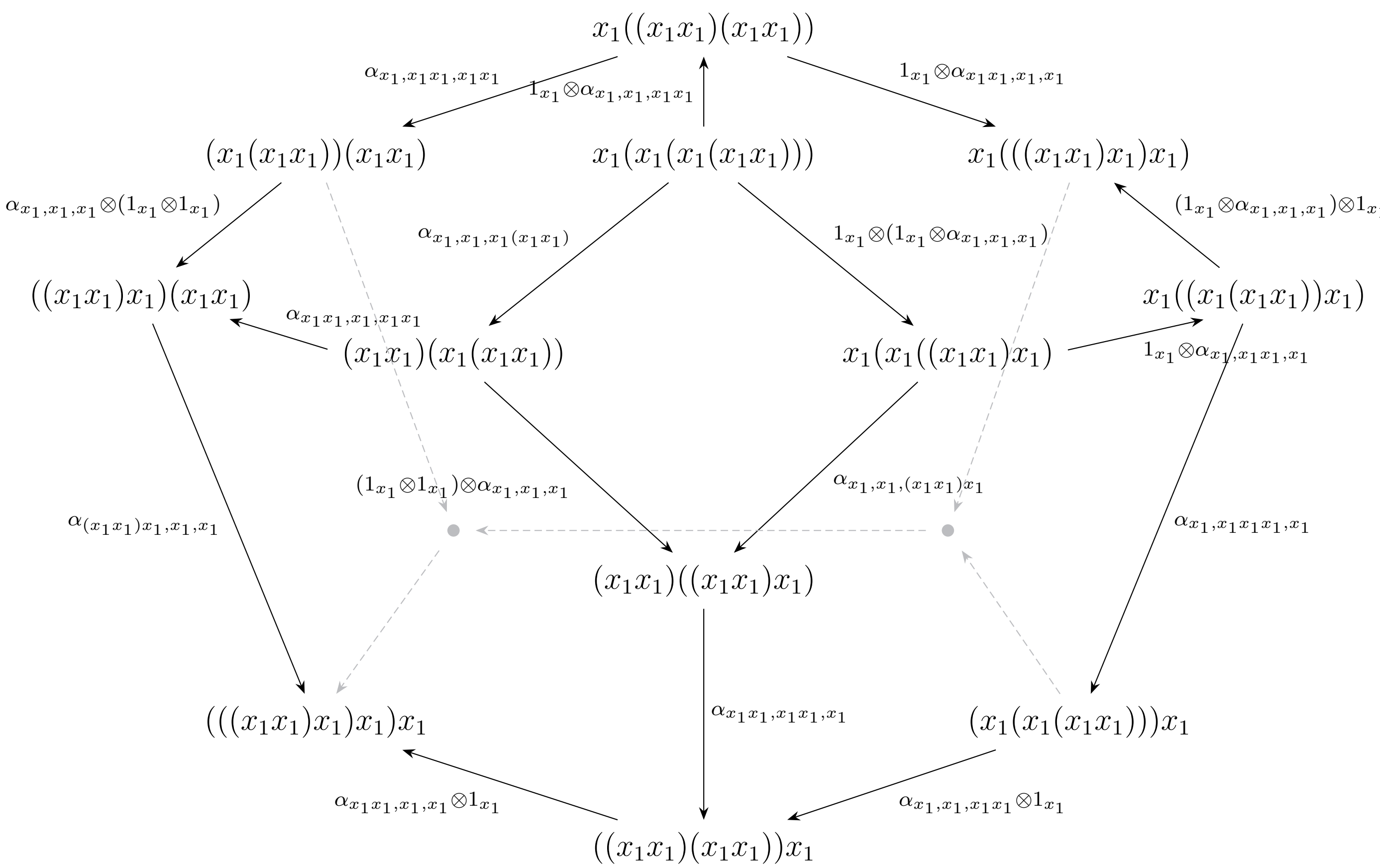
So far, our functor makes sense. Moreover, we already knew that the above pentagon commutes for all $A, B, C, D \in \mathcal{M}$. Thus, what about diagram **??**?

Again, functors preserve diagrams. Therefore, the commutative diagram in $\mathcal{W}$ (see next page) is mapped by $\Phi_{\mathrm{id}}$ to the commutative diagram of natural transformations in $\mathbf{It}(\mathcal{M})$ between functors (see second page) and as functors from $\mathcal{M}^5 \longrightarrow \mathcal{M}$, we may substitute any $A, B, C, D, E$ to obtain the commutative diagram in $\mathcal{M}$ (on the third page).

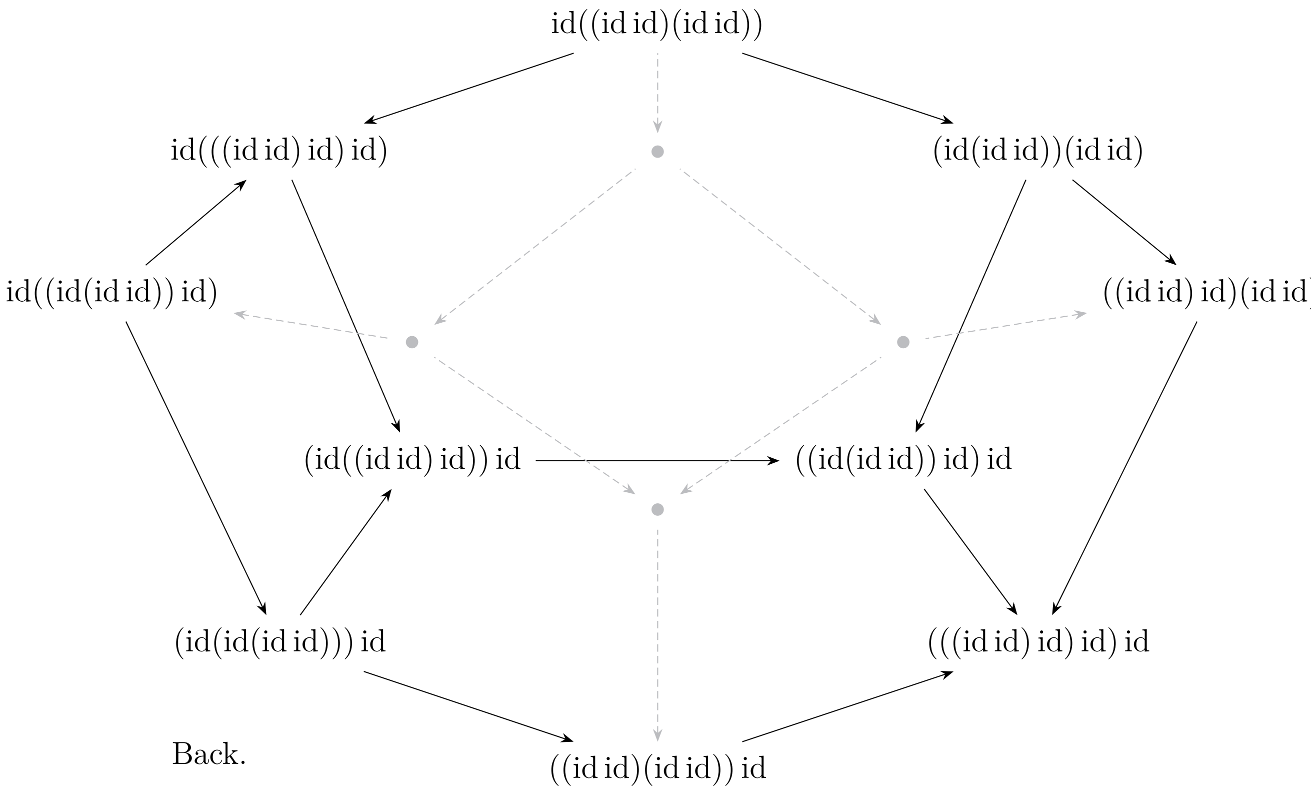
Front. (Note that the product $\otimes$ in $\mathcal{W}$ has been suppressed).



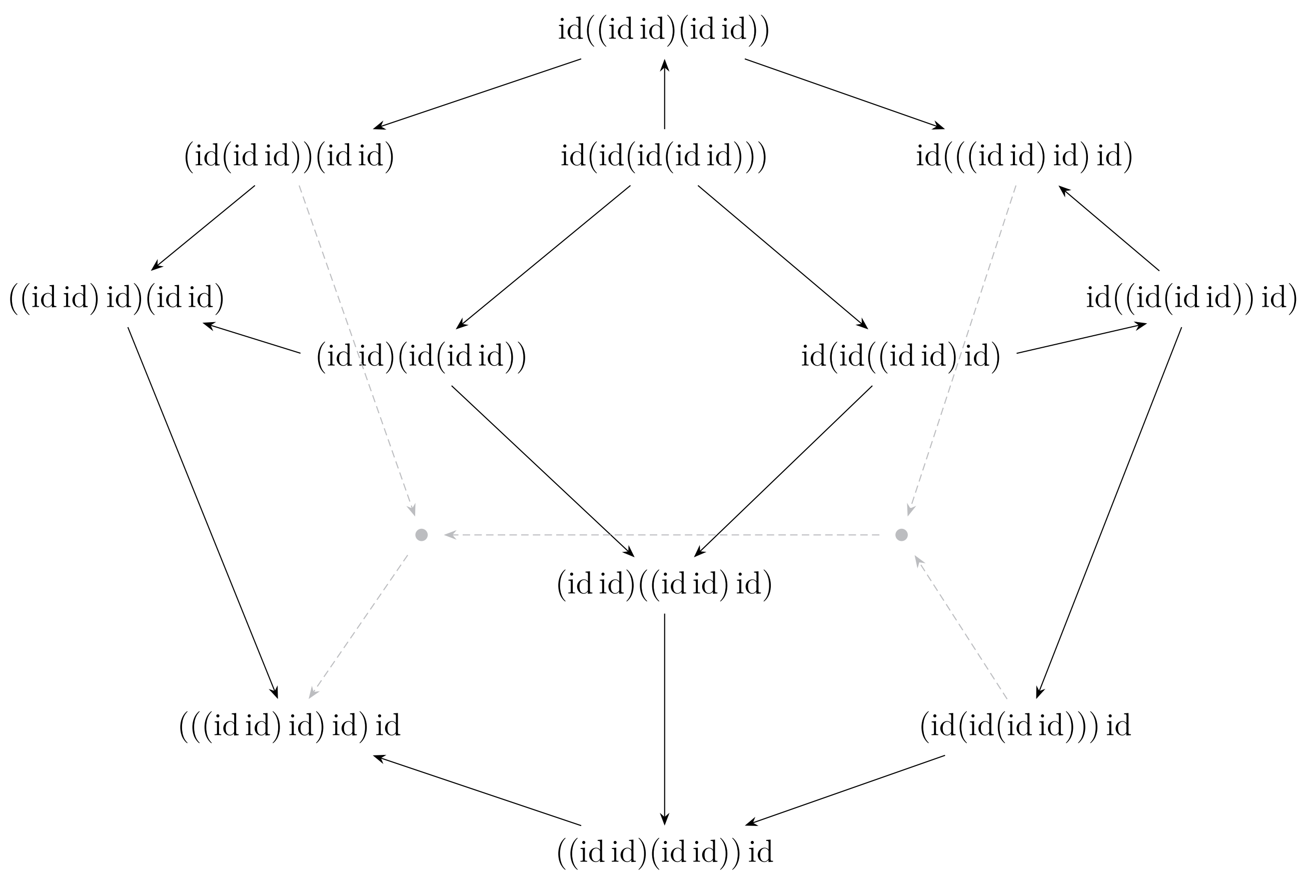

Back.



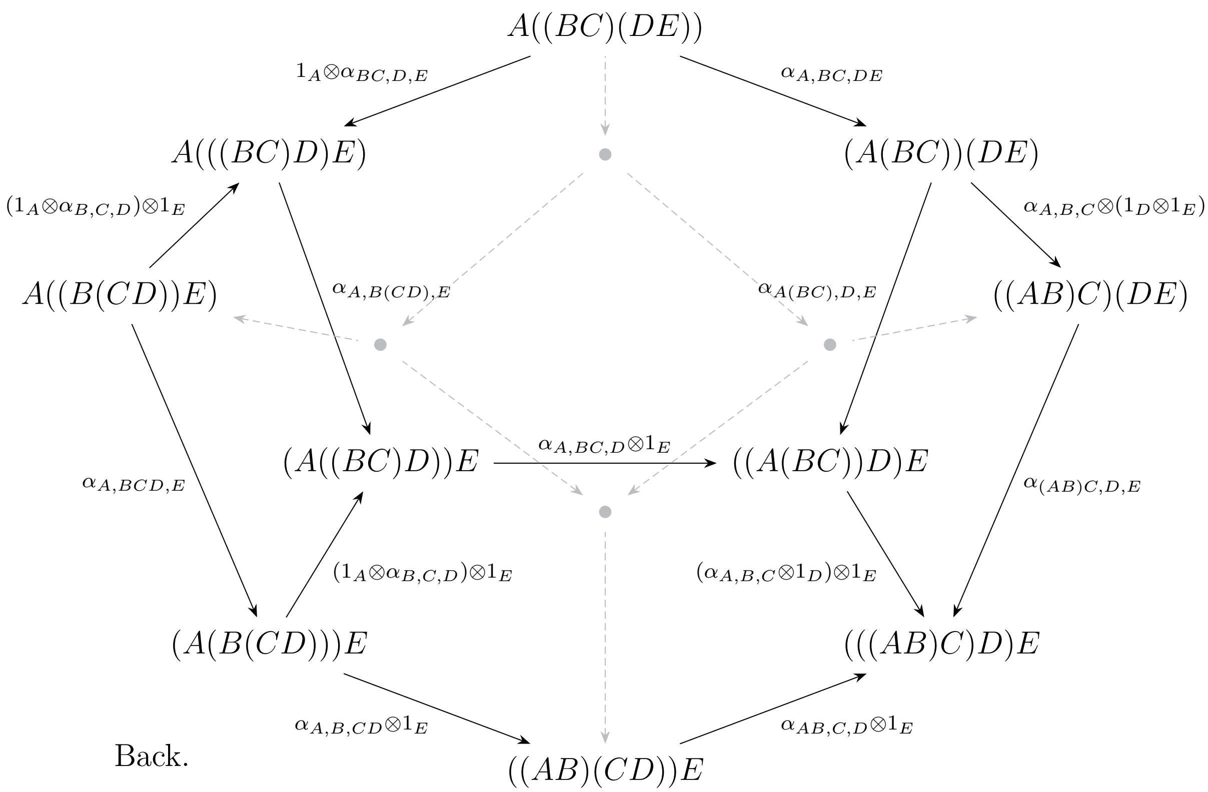
Front. (Note that the product $\odot$ and the associators in $\mathbf{It}(\mathcal{M})$ are suppressed.)



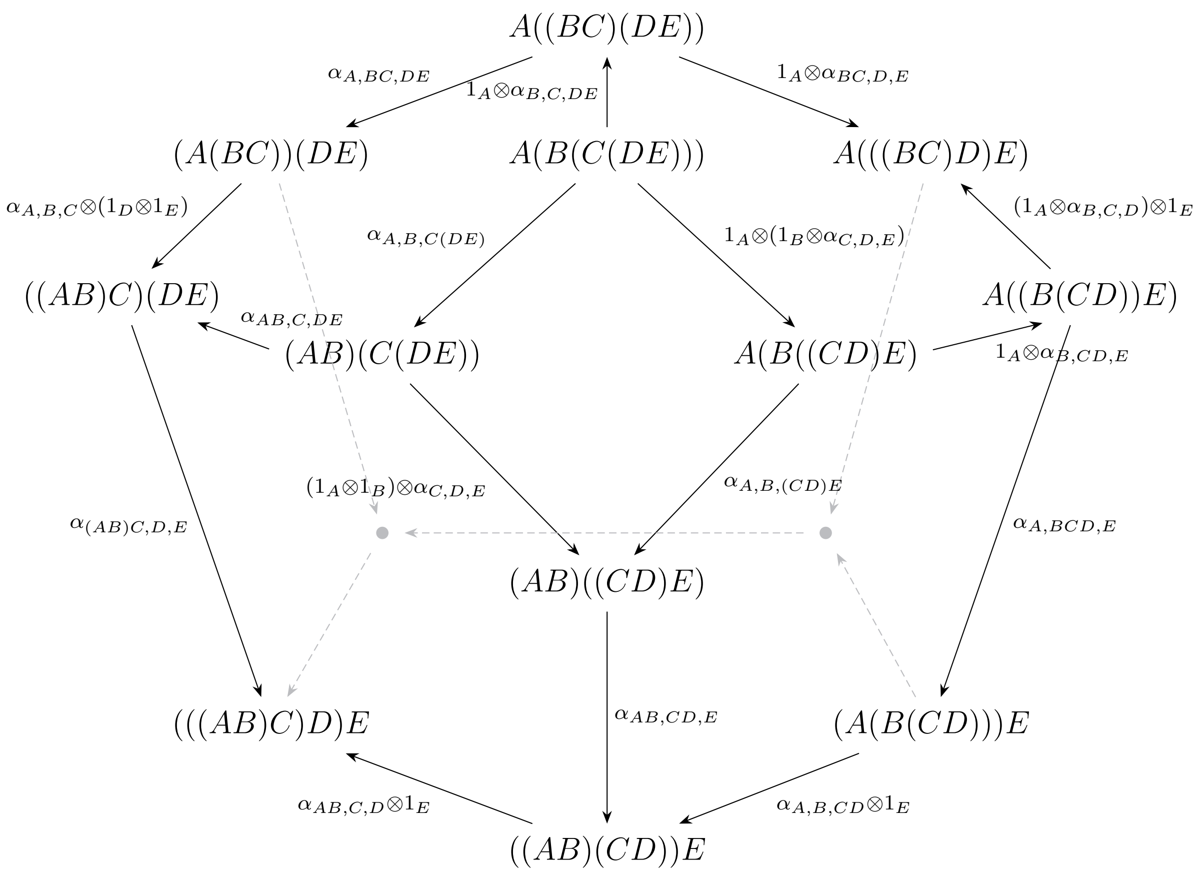

Back.



Front. (Note that the product $\otimes$ in $\mathcal{M}$ is suppressed.)




Back.



This process continues for every possible diagram in $\mathcal{W}$. Each diagram in $\mathcal{W}$ is mapped to a corresponding diagram in $\mathbf{It}(\mathcal{M})$ made up of identity functors, and with the identity functor, we are free to substitute whatever instance of $A, B, C, \cdots \in M$ in it. The arrows between the identity functors are natural transformations which reduce to instances of $\alpha, \rho, \lambda$ in $\mathcal{M}$ upon substituting objects in the identity functor. What matters here is the functoriality of $\Phi_I$. It guarantees that all the diagrams obtained as the image of $\Phi_{\text{id}}$ will commute.

This completes our work towards proving Mac Lane's Coherence Theorem.

## 7.5 Braided and Symmetric Monoidal Categories

Braided and symmetric monoidal categories serve as some of the most fruitful and most studied environments of monoidal categories. The formulation of these categories may seem mysterious and random, but they have been recognized as important in their applications to physics. Specifically, braided monoidal categories were first defined by Joyal-Street in an attempt to abstract the solutions to the Yang-Baxter equation, an important equation of matrices in statistical mechanics. It turns out that braided monoidal categories are exactly the categorical environment one needs to describe the category of representations of a Hopf algebra $\mathbf{Rep}(H)$. This then allows us a machine which produces solutions to the Yang-Baxter equation, ultimately letting us describe families of such solutions. But it gets even more interesting: the Yang-Baxter equation turns out to be the necessary criteria to establish a representation of the Braid group; such a representation is a knot invariant, so this is something of interest to both mathematicians and physicists.

Before we dive into what exactly braided monoidal categories are, we'll introduce the concept of braids.

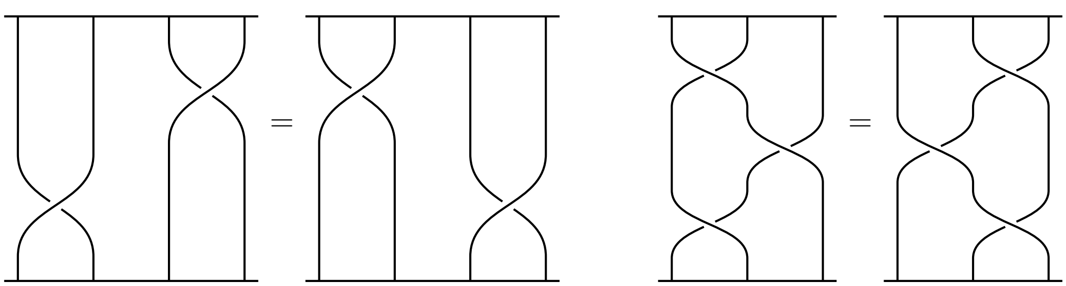
**Definition 7.5.1.** The $n$-th braid group $B_n$ consists of braids on $n$-strands whose group product is braid composition. More rigorously,

$$B_n = \left\langle \sigma_1, \dots, \sigma_n, \sigma_1^{-1}, \dots, \sigma_n^{-1} \mid (1), (2) \right\rangle$$

where $(1), (2)$ are generator relations described below.

1. $\sigma_i\sigma_j = \sigma_j\sigma_i$ whenever $|i-j| > 1$
2. $\sigma_{i+1}\sigma_i\sigma_{i+1} = \sigma_i\sigma_{i+1}\sigma_i$.

Relations (1) and (2) are imposed in order to reflect physical reality. Below the relations are pictured on a three strands.



*Geometrically, the above braids are clearly equivalent. Algebraically this translates to the statements $\sigma_3\sigma_1 = \sigma_1\sigma_3$ and $\sigma_1\sigma_2\sigma_1 = \sigma_2\sigma_1\sigma_2$.*

The first two braids represent $\sigma_3\sigma_1$ and $\sigma_1\sigma_3$. Clearly, these are physically equal. Note however this would not work if they were adjacent, i.e., $\sigma_2\sigma_1 \neq \sigma_2\sigma_1$. Hence we set $\sigma_i\sigma_j = \sigma_j\sigma_i$ for $|i-j| > 1$. For the second pair of braids, it may take some staring to see that they are physically equal. As we shall see, the relation $\sigma_{i+1}\sigma_i\sigma_{i+1} = \sigma_i\sigma_{i+1}\sigma_i$, called the **braid relation**, is of deep importance.

**Definition 7.5.2.** A **Braided Monoidal Category** $\mathcal{C}$ is a monoidal category $(\mathcal{C}, \otimes, I)$ equipped additionally with a natural transformation, know as the "twist" morphism

$$\sigma_{A,B} : A \otimes B \longrightarrow B \otimes A \qquad \textbf{(Twist Morphism)}$$

such that the following diagrams commute for all objects $A, B, C$ of $\mathcal{C}$.

$$\begin{tikzcd}
 & X \otimes (Y \otimes Z) \arrow[r, "\sigma_{X,Y\otimes Z}"] & (Y \otimes Z) \otimes X \arrow[dr, "\alpha_{Y,Z,X}"] & \\
(X \otimes Y) \otimes Z \arrow[ur, "\alpha_{X,Y,Z}"] \arrow[dr, "\sigma_{X,Y}\otimes 1_Z"'] & & & Y \otimes (Z \otimes X) \\
 & (Y \otimes X) \otimes Z \arrow[r, "\alpha_{Y,X,Z}"'] & Y \otimes (X \otimes Z) \arrow[ur, "1_Y \otimes \sigma_{X,Z}"'] &
\end{tikzcd} \tag{7.3}$$

$$\begin{tikzcd}
 & (X \otimes Y) \otimes Z \arrow[r, "\sigma_{X\otimes Y,Z}"] & Z \otimes (X \otimes Y) \arrow[dr, "\alpha^{-1}_{Y,Z,X}"] & \\
X \otimes (Y \otimes Z) \arrow[ur, "\alpha^{-1}_{X,Y,Z}"] \arrow[dr, "1_X\otimes \sigma_{Y,Z}"'] & & & (Z \otimes X) \otimes Y \\
 & X \otimes (Z \otimes Y) \arrow[r, "\alpha^{-1}_{X,Z,Y}"'] & (X \otimes Z) \otimes Y \arrow[ur, "\sigma_{X,Z} \otimes 1_Y"'] &
\end{tikzcd} \tag{7.4}$$

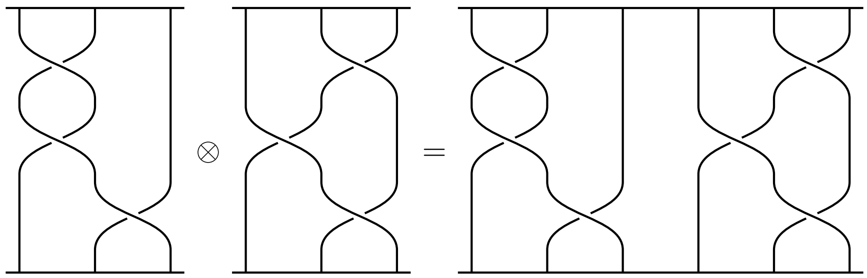
Note that just because we have a twist morphism, it is not necessarily the case that $\sigma_{B,A} \circ \sigma_{A,B} = 1_{A\otimes B}$. That is, applying the twist morphism twice is not guaranteed to give you back the identity. This case is treated separately.

---

**Example 7.5.3.** The canonical example of a braided monoidal category is the braid category $\mathbb{B}$. This is the category where:
**Objects.** All integers $n \geq 0$.
**Morphisms.** For any two integers $m, n$, we have that

$$\text{Hom}_{\mathbb{B}}(n, m) = \begin{cases} B_n & \text{if } n = m \\ \varnothing & \text{if } n \neq m \end{cases}$$

Composition in this category is simply braid composition. We can introduce a tensor product $\otimes$ on $\mathbb{B}$ where on objects $n \otimes m = n + m$ while on morphisms $\alpha \otimes \beta$ is the direct sum braid. The direct sum braid is simply the braid obtained by placing two braids side-by-side.

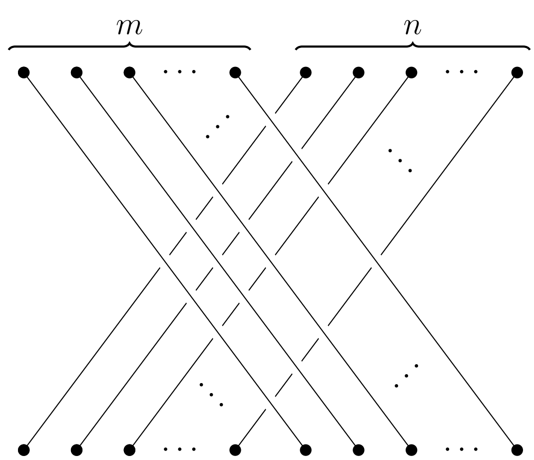


*The braids $\sigma_1\sigma_1\sigma_2$ and $\sigma_2\sigma_1\sigma_2$ are summed together to obtain the braid $\sigma_1\sigma_1\sigma_2\sigma_5\sigma_4\sigma_5$ above on the right.*

With an identity object being the empty braid, we see that $\mathbb{B}$ is a strict monoidal category. The associators and unitors are simply identity morphisms. However, this category also have a natural braiding structure. For any two objects $n, m$, introduce the braiding

$$\sigma_{n,m} : n + m \xrightarrow{\sim} m + n$$

where on objects the addition is simply permuted; on morphisms, however, $\sigma_{n,m}$ corresponds to the braid of length $n + m$ as below.




It is a simple exercise to show that this satisfies the hexagon axioms; the task is simplified due to the fact that the associators are identities. While this category may seem like a boring example, it plays a critical role in demonstrating coherence for braided monoidal categories, something we will do later.

---

**Example 7.5.4.** Let $\mathbf{GrMod}_R$ be the category of graded $R$-modules $M = \{M_n\}_{n=1}^{\infty}$. Recall from 7.1 That $\mathbf{GrMod}_R$ forms a monoidal category. The tensor product of two graded $R$ modules $M = \{M_n\}_{n=1}^{\infty}$ and $P = \{P_n\}_{n=1}^{\infty}$ is the graded $R$-module $M \otimes P$ whose $n$-th level is given by

$$(M \otimes P)_n = \bigoplus_{i+j=n} M_i \otimes P_j.$$

We can additionally introduce a braiding on this category for each invertible elements $k \in R$; specifically, we define the braiding $\sigma_{M,P} : M \otimes P \longrightarrow P \otimes M$ to be the graded module homomorphism whose $n$-th degree is

$$
\begin{aligned}
(\sigma_{M,P})_n : \bigoplus_{i+j=n} M_i \otimes P_j &\longrightarrow \bigoplus_{i+j=n} P_j \otimes M_i \\
(m \otimes p) &\longmapsto k^{ij} p \otimes m
\end{aligned}
$$

whenever $m \in M_i$ and $p \in P_j$. Observe that with this braiding we get that

$$
\begin{tikzcd}
 & m \otimes (p \otimes q) \arrow[r, mapsto] & r^{(j+k)i}(p \otimes q) \otimes m \arrow[dr, mapsto] & \\
 & & & r^{(j+k)i} p \otimes (q \otimes m) \\
(m \otimes p) \otimes q \arrow[uur, mapsto] \arrow[ddr, mapsto] & & & = \\
 & & & r^{ij} p \otimes (r^{ik} m \otimes q) \\
 & r^{ij}(p \otimes m) \otimes q \arrow[r, mapsto] & r^{ij} p \otimes (m \otimes q) \arrow[ur, mapsto] &
\end{tikzcd}
$$

which clearly commutes. The second hexagon axiom is also easily seen to be satisfied:

$$
\begin{tikzcd}
 & (m \otimes p) \otimes q \arrow[r, mapsto] & r^{(i+j)k} q \otimes (m \otimes p) \arrow[dr, mapsto] & \\
 & & & r^{(i+j)k}(p \otimes m) \otimes n \\
m \otimes (p \otimes q) \arrow[uur, mapsto] \arrow[ddr, mapsto] & & & = \\
 & & & r^{jk}(r^{ik} p \otimes m) \otimes n \\
 & r^{jk} m \otimes (q \otimes p) \arrow[r, mapsto] & r^{jk}(m \otimes q) \otimes p \arrow[ur, mapsto] &
\end{tikzcd}
$$

Thus we see that $\mathbf{GrMod}_R$ is more than just a monoidal category; each invertible element of $R$ induces a braiding, making it a braided monoidal category as well.

---

**Example 7.5.5.** If $M$ is monoidal, we can recall from Example 7.1 that the functor category $\mathcal{C}^M$ is also monoidal. If additionally we have that $M$ is braided with a braiding $\sigma_{A,B} : A \otimes B \longrightarrow B \otimes A$, then we can extend this to a braiding on the functor category of $\mathcal{C}^M$ by defining, for two functors $F, G : \mathcal{C} \longrightarrow M$, the natural transformation

$$
\beta_{F,G} : F \otimes G \longrightarrow G \otimes F
$$

defined pointwise for each $A \in \mathcal{C}$ as the morphism

$$
(\beta_{F,G})_A = \sigma_{F(A),G(A)} : F(A) \otimes G(A) \xrightarrow{\sim} G(A) \otimes F(A).
$$

One can then check that this natural transformation satisfies the braided hexagon axioms since the braiding $\sigma$ in $M$ does, so that $\mathcal{C}^M$ is additionally braided if $M$ is additionally braided.

**Definition 7.5.6.** A **Symmetric Monoidal Category** $\mathcal{C}$ is a braided monoidal category such that, for the twist morphism,

$$\sigma_{B,A} \circ \sigma_{A,B} = 1_{A \otimes B}.$$

Symmetric monoidal categories are basically monoidal categories which collapse the information which braided monoidal categories have the potential to encode. Their environment is much simpler, but at the cost of information.

**Example 7.5.7.** Recall from the previous examples that $\mathbf{GrMod}_R$ can be treated as a braided monoidal category. A braiding is given an invertible element $r \in R$. However, consider the idempotent elements of this ring, i.e., the elements $r \in R$ such that $r^2 = 1$. Then we see that these elements not only give rise to braidings

$$(\sigma_{M,P})_n : \bigoplus_{i+j=n} M_i \otimes P_j \longrightarrow \bigoplus_{i+j=n} P_j \otimes M_i$$
$$(m \otimes p) \longmapsto k^{ij} p \otimes m$$

but these braidings have the property that $\sigma_{M,P} \circ \sigma_{P,M} = 1_{M \otimes P}$, since $r = 1$. Hence the category of graded modules may be specially treated as symmetric monoidal categories whenever there is an idempotent element of the ring $R$.

**Example 7.5.8.** Recall from 7.1 that the permutation category $\mathbb{P}$ forms a monoidal category where objects are nonnegative integers and homsets are given by the symmetric groups. The monoidal product $\otimes$ simply sums the object, while two permutations $\tau \in S_n$ and $\rho \in S_m$ are sent to the direct sum permutation $\tau \otimes \rho \in S_{n+m}$ (this permutation simply horizontally stacks).

In this category, we can introduce a symmetric braiding $\sigma_{n,m} : n + m \longrightarrow m + n$ to be the unique permutation $\sigma_{n,m} \in S_{n+m}$ pictured below.

$$\begin{array}{c}(1, 2, \dots, n, n+1, n+2, \dots, n+m)\\ \Big\downarrow \sigma_{n,m} \\ (n+1, n+2, \dots, n+m, 1, 2, \dots, n)\end{array}$$

One thing to notice is that this is the underlying permutation of braid given in Figure 7.5. With the existence of this element of $S_{n+m}$ for every pair of objects $n, m$ in $\mathbb{P}$, we see that the permutation category is actually symmetric monoidal.

**Definition 7.5.9.** A **PROP**, an acronym coined by Mac Lane for "Product and Permutation Category", is a symmetric monoidal category $\mathbb{P}$ containing the category $(\mathbb{N}, 0, +)$.

---

**Example 7.5.10.** Consider the category **FinSet**, where the objects are natural numbers $n$ and a morphism $f : n \longrightarrow m$ is a function from a set of size $n$ to one of size $m$.

Here, we necessarily include 0 as an object; this denotes the empty set. First we demonstrate that this is monoidal. Let $n, m$ be any integers. Then we'll show that $+ : \textbf{FinSet} \times \textbf{Finset} \longrightarrow \textbf{FinSet}$ is a bifunctor. First, we acknowledge that $n + m \in \textbf{FinSet}$.

Next, consider the set of morphisms

$$\begin{aligned} h &: k \longrightarrow n \qquad & f &: n \longrightarrow n' \\ j &: l \longrightarrow m \qquad & g &: m \longrightarrow m'. \end{aligned}$$

Let $S_k$ be the set of $k$ elements. Now since $f, g$ are functions, we know that $f : S_n \longrightarrow S_{n'}$ and $g : S_m \longrightarrow S_{m'}$ for some sets in **Set**.Then we can define $f + g : (n + n') \longrightarrow (m + m')$ to be the function in **Set** where

$$f + g : S_n \amalg S_{n'} \longrightarrow S_m \amalg S_{m'}.$$

where

$$(f+g)(x,i) = \begin{cases} (f(x),0) & \text{if } i = 0 \\ (g(x),1) & \text{if } i = 1. \end{cases}$$

Hence $f + g$ makes sense in **FinSet** as morphism $f + g : (n + n') \longrightarrow (m + m')$.

Now consider the morphisms $f \circ h$ and $g \circ j$. Observe that $f \circ h + g \circ j : k + l \longrightarrow n' + m'$. This is then the function

$$f \circ h + g \circ j : S_k \amalg S_l \longrightarrow S_{n'} \amalg S_{m'}$$

but note that

$$f \circ h + g \circ j : S_k \amalg S_l \longrightarrow S_{n'} \amalg S_{m'} = (f + g) \circ (h + j)$$

Hence we must have that $(f + g) \circ (h + j) = f \circ h + g \circ j$, so that we have that $+$ is a bifunctor.

Now we show that this is a monoidal category. Define the natural isomorphisms

$$\begin{aligned} \alpha_{n,m,p} &: n + (m + p) \xrightarrow{\sim} (n + m) + p \\ \lambda_n &: 0 + n \xrightarrow{\sim} n \\ \rho_n &: n + 0 \xrightarrow{\sim} n. \end{aligned}$$

We can describe these functions in further detail. Observe that $\alpha_{n,m,p}$ can be realized to be a function where

$$\alpha_{n,m,p} : S_n \amalg (S_m \amalg S_p) \xrightarrow{\sim} (S_n \amalg S_m) \amalg S_p.$$

Elements of $S_n \amalg (S_m \amalg S_p)$ will be either $(x, 0)$ where $x \in S_n$, or $(x, 1)$ where $x \in S_m \amalg S_p$. In turn, the elements of this set are of the form $(y, 0)$ where $y \in S_m$ and $(y, 1)$ where $y \in S_p$.

On the other hand, elements of $(S_n \amalg S_m) \amalg S_p$ are of the form $(x', 0)$ if $x' \in S_n \amalg S_m$ or are of the form $(x', 1)$ if $x' \in S_p$. Furthermore, elements of $S_n \amalg S_m$ are of the form $(y', 0)$ if $y' \in S_n$

and $(y', 1)$ if $y' \in S_m$.

Now we can explicitly define $\alpha_{n,m,p}$ as

$$\alpha_{n,m,p}(x,i) = \begin{cases} ((x,0),0) & \text{if } i = 0 \\ ((y,1),0) & \text{if } i = 1 \text{ and } x = (y,0) \\ (y,1) & \text{if } i = 1 \text{ and } x = (y,1) \end{cases} \tag{7.5}$$

and $\lambda$ as

$$\lambda_n(x,1) = x$$

and $\rho$ as

$$\rho_n(x,0) = x.$$

Note for both $\lambda$ and $\rho$, there is only one case for $(x,i)$ since for $\lambda$, $i$ is never 0 and for $\rho$, $i$ is never 1.

All of these establish a bijection, and hence an isomorphism. Now to demonstrate that they are natural, consider $f : n \longrightarrow n'$, $g : m \longrightarrow m'$ and $h : p \longrightarrow p'$. First, we'll want to show that the diagram

$$\begin{tikzcd} n + (m+p) \arrow[r, "\alpha_{n,m,p}"] \arrow[d, "f+(g+h)"'] & (n+m)+p \arrow[d, "(f+g)+h"] \\ n' + (m'+p') \arrow[r, "\alpha_{n',m',p'}"'] & (n'+m')+p' \end{tikzcd}$$

commutes, which we can do by a case-by-case basis. First we follow the path

$$[(f+g)+h] \circ \alpha_{n,m,p} : S_n \amalg (S_m \amalg S_p) \longrightarrow (S_{n'} \amalg S_{m'}) \amalg S_{p'}.$$

and then show it is equivalent to the other path.

$\boldsymbol{i = 0}$ If the input is $(x,0)$, we see that $\alpha_{n,m,p}(x,i) = ((x,0),0)$. If this is fed into $(f+g)+h$, the output will be $(f+g)(x,0)$, whose output will be $((f(x),0),0)$.
However, suppose we first put $(x,0)$ into $f + (g+h)$. Then we would have directly obtain $(f(x),0)$. Feeding this into $\alpha_{n',m',p'}$, we would get $((f(x),0),0)$. Hence we obtain naturality in this case.

$\boldsymbol{i = 1.}$ Suppose now the input is $(x,1)$. Then either $x = (y,0)$ with $y \in S_m$ or $(y,1)$ where $y \in S_p$.

$\boldsymbol{y \in S_m.}$ Suppose $x = (y,0)$. Then we see that $\alpha_{n,m,p}(x,1) = ((y,1),0)$. Plugging this into $(f+g)+h$, we get

$$[(f+g)+h]((y,1),0) = ([f+g](y,1),0) = ((g(y),1),0).$$

However, we also could have obtained this value by first starting with $f + (g + h)$. In this case,

$$[f + (g + h)]((y, 0), 1) = ([g + h](y, 0), 1) = ((g(y), 0), 1).$$

Plugging this into $\alpha_{n',m',p'}$, we then get that

$$\alpha_{n',m',p'}((g(y), 0), 1) = ((g(y), 1), 0).$$

Hence the two paths are equivalent.

***$y \in S_p$.*** Suppose $x = (y, 1)$, Then we have that $\alpha_{n,m,p}((y, 1), 1) = (y, 1)$. Sending this into $(f + g) + h$, we get

$$[(f + g) + h](y, 1) = (h(y), 1).$$

However, we could have achieved this value by first plugging $((y, 1), 1)$ into $f+(g+h)$:

$$[f + (g + h)]((y, 1), 1) = ([g + h](y, 1), 1) = ((h(y), 1), 1).$$

Then sending this into $\alpha_{n',m',p'}$, we get

$$\alpha_{n',m',p'}((h(y), 1), 1) = (h(y), 1).$$

Thus the two paths are equivalent.

Hence we see that this diagram does commute, so that $\alpha$ is natural.

[Show naturality works for $\lambda$ and $\rho$.]

Now we show that these natural isomorphisms satisfy the monoidal properties. Specifically, we'll show that the diagram

$$\begin{tikzcd} n + (0 + m) \arrow[rr, "\alpha_{n,0,m}"] \arrow[dr, "1_n + \lambda_m"'] & & (n + 0) + m \arrow[dl, "\rho_n + 1_m"] \\ & n + m & \end{tikzcd} .$$

must commute. To do this, we consider how these functions are realized in **Set**. If we consider $(x, i) \in S_n \amalg (\varnothing \amalg S_m)$, we see that we have two cases to consider.

***$i = 0$.*** If $i = 0$, then we see that $\alpha_{n,0,m}(x, 0) = ((x, 0), 0)$. Sending this into $\rho_n + 1_m$, we get that $[\rho_m + 1_m]((x, 0), 0) = (\rho(x, 0), 0) = (x, 0)$.

On the other hand, we could obtain this value by directly sending $(x, 0)$ into $1_n + \lambda_m$. Observe that $[1_n + \lambda_m](x, 0) = (1_n(x), 0) = (x, 0)$. Hence the diagram commutes for this case.

***$i = 1$.*** If $i = 1$, then our element is of the form $(x, 1)$. However, we know that $x = ((x, 1), 0)$, since $(x, 1) \in 0 + m$. Thus observe that $\alpha_{n,0,m}((x, 1), 1) = (x, 1)$. Consequently, we get that $[\rho_n + 1_m](x, 1) = (1_m(x), 1) = (x, 1)$.

On the other hand, we can start instead be evaluating $[1_n + \lambda_m]((x,1),1) = (\lambda(x,1),1) = (x,1)$. Hence the diagram commutes in this case.

Thus we see that this diagram holds for all naturals $n, m$.

# 7.6 Coherence for Braided Monoidal Categories

We saw with monoidal categories that ultimately everything we were saying *made sense*. That is, we saw that our definition does not give us an contradictions, and that we can obtain a significant coherence result which ultimately allows us to not worry about the particular parenthesization of a monoidal product. Further, we saw that diagrams freely built from associators and unitors were all commutative.

With braided monoidal categories we can get a similar statement. This time, however, it is a bit weaker, although it is nevertheless extremely useful. It was Joyal and Street in the 1993 paper who both first proved the coherence for braided monoidal categories. Their work heavily relies on the work of G.M. Kelly, and they use very slick, higher categorical tricks.

In this section, we spell out those tricks.

**Definition 7.6.1** (Joyal-Street)**.** Let $\mathcal{A}$ be a category with $\mathcal{V}$ a monoidal category. Suppose $T : \mathcal{A} \longrightarrow \mathcal{V}$ is a functor. We define a **Yang-Baxter operator** to be a family of isomorphisms

$$y_{A,B} : T(A) \otimes T(B) \xrightarrow{\sim} T(B) \otimes T(A).$$

for each $A, B \in \mathcal{A}$ such that the diagram below commutes. such that the diagram below commutes.

$$\begin{tikzcd}
& T(A)\otimes T(B)\otimes T(C) \arrow[r, "y_{A,B}\otimes 1"] \arrow[dl, "1\otimes B_{B,C}"'] & T(B)\otimes T(A)\otimes T(C) \arrow[dr, "1\otimes y_{A,C}"] & \\
T(A)\otimes T(C)\otimes T(B) \arrow[dr, "y_{A,C}\otimes 1"'] & & & T(B)\otimes T(C)\otimes T(A) \arrow[dl, "y_{B,C}\otimes 1"] \\
& T(C)\otimes T(A)\otimes T(B) \arrow[r, "1\otimes y_{A,B}"'] & T(C)\otimes T(B)\otimes T(A) &
\end{tikzcd}$$

Note that here we omit the associators although they are implicitly included in the diagram. Note also that, for any functor $T : \mathcal{A} \longrightarrow \mathcal{V}$ with $\mathcal{V}$ a braided monoidal category, $T$ trivially has a Yang-Baxter operator $y$ where we set

$$y_{A,B} = \sigma_{T(A),T(B)}.$$

Before we move forward we introduce a notion that can be found in [**?**], originally from [**?**]. For our purposes, we will denote the category obtained via disjoint unions of the symmetric groups $S_n$ as $\mathbb{P}$. That is, the objects of $\mathbb{P}$ are natural numbers and

$$\mathrm{Hom}_{\mathbb{P}}(n, m) = \begin{cases} S_n & \text{if } n = m \\ \varnothing & \text{if } n \neq m \end{cases}$$

**Definition 7.6.2.** Let $\mathcal{A}$ be a category and suppose suppose $\mathcal{D} \in \mathbf{Cat}/\mathbb{P}$. That is, $\mathcal{D}$ is a category with an associated functor $\Gamma : \mathcal{D} \longrightarrow \mathbb{P}$. Then we define the category $\mathcal{D} \int \mathcal{A}$ where
**Objects.** Finite strings $[A_1, A_2, \dots, A_n]$ with $A_i \in \mathcal{A}$
**Morphisms.** For two strings $[A_1, \dots, A_n]$ and $[B_1, \dots, B_n]$, denoted as $[A_i]$ and $[B_i]$,

$$\text{Hom}_{\mathcal{D} \int \mathcal{A}}\Big([A_i], [B_i]\Big) = \Big\{(\alpha, f_1, \dots, f_n) \mid f_i \in \text{Hom}_{\mathcal{A}}(A_i, B_{\sigma(i)})\Big\}$$

Here $\alpha$ is a morphism of $\mathcal{D}$ such that $\Gamma(\alpha) = \sigma \in S_n$. Finally, we allow no morphisms between two different strings of different length.

For any category $\mathcal{A}$, there exists a natural inclusion functor

$$i_{\mathcal{A}} : \mathcal{A} \longrightarrow \mathcal{D} \int \mathcal{A}$$
$$i_{\mathcal{A}}(A) = [A] \qquad i_{\mathcal{A}}(f : A \longrightarrow B) = (e_1, f) : [A] \longrightarrow [B]$$

where $e_1$ is the sole element of $S_1$. This functor will be useful for us later. Next we formalize the following category which can be thought of as a generalized functor category.

**Definition 7.6.3.** Let $\mathcal{A}, \mathcal{B}$ be categories. Denote the category $\{\mathcal{A}, \mathcal{B}\}$ as the category with objects $(n, F : \mathcal{A}^n \longrightarrow \mathcal{B})$ whose morphisms are

$$\text{Hom}_{\{\mathcal{A},\mathcal{B}\}}\Big((n, T), (m, S)\Big) = \begin{cases} \{(\sigma, \eta : \sigma \cdot T \longrightarrow S)\} & \text{if } n = m \\ \varnothing & \text{if } n \neq m. \end{cases}$$

Here $\sigma \in S_n$, and $\eta : \sigma \cdot T \longrightarrow S$ is a natural transformation from the functor $\sigma \cdot T$ defined pointwise as

$$\sigma \cdot T(A_1, A_2, \dots, A_n) = T(A_{\sigma(1)}, \dots, A_{\sigma(n)})$$

to the functor $S$.

There are two things we need to say about this category. First, for any generalized functor category $\{\mathcal{A}, \mathcal{V}\}$, there exists a a projection functor $\Gamma : \{\mathcal{A}, \mathcal{B}\} \longrightarrow \mathbb{P}$ defined on objects and morphisms as

$$\Gamma(n, T : \mathcal{A}^n \longrightarrow \mathcal{B}) = n \qquad \Gamma(\sigma, \eta : \sigma \cdot T \longrightarrow S) = \sigma.$$

Hence we see that each category $\{\mathcal{A}, \mathcal{B}\}$ is actually a member of $\mathbf{Cat}/\mathbb{P}$, because it always comes equipped with a functor into $\mathbb{P}$.

Second, if $\mathcal{V}$ is a strict monoidal category, then so is $\{\mathcal{A}, \mathcal{V}\}$. One can see this by defining for two functors $T : \mathcal{A}^n \longrightarrow \mathcal{V}$ and $S : \mathcal{A}^m \longrightarrow \mathcal{V}$ the functor $T \otimes S : \mathcal{A}^{n+m} \longrightarrow \mathcal{V}$ which is a functor that can be defined pointwise as

$$(T \otimes S)(A_1, \dots, A_{n+m}) = T(A_1, \dots, A_n) \otimes S(A_{n+1}, \dots, A_{n+m}).$$

Thus if $\mathcal{V}$ is strict, then so it $\{\mathcal{A}, \mathcal{V}\}$.

What is useful about this construction is that Kelly showed that the functors

$$(-)\int A : \mathbf{Cat}/\mathbb{P} \longrightarrow \mathbf{Cat} \qquad \{A,(-)\} : \mathbf{Cat} \longrightarrow \mathbf{Cat}/\mathbb{P}$$

form an adjunction. We use this in the next proposition, which is also aided by the following lemma.

**Lemma 7.6.4.** Let $\mathcal{V}$ be a strict monoidal category. Suppose $T : \mathbf{1} \longrightarrow \mathcal{V}$ has a Yang-Baxter operator $y$. Then there exists a unique strict monoidal functor $T' : \mathbb{B} \longrightarrow \mathcal{V}$ such that the diagram below commutes.

$$\begin{tikzcd}
\mathbf{1} \arrow[r, "i"] \arrow[dr, "T"'] & \mathbb{B} \arrow[d, dashed, "T'"] \\
& \mathcal{V}
\end{tikzcd}$$

Further, we have that $T'(\sigma) = y$.

*Proof.* Denote the element of $\mathbf{1}$ as $\bullet$. Then $T(\bullet) = X$ for some $X \in \mathcal{V}$. Towards a definition of $T'$, let $T' : \mathbb{B} \longrightarrow \mathcal{V}$ be defined on objects as $T'(1) = X$. If we force $T'$ to be strict, this will define its value on all objects of $\mathbb{B}$. On morphisms, first observe that each $\beta \in B_n$ can be expressed in terms of its generators $\sigma_i$. Hence it suffices to define the action of $T'$ on a generator $\sigma_i$, and we do this naturally as:

$$T'(\sigma_i) = 1_X^{\otimes(i-1)} \otimes y_{X,X} \otimes 1_X^{\otimes(n-i-1)} : X^{\otimes n} \longrightarrow X^{\otimes n}$$

We then define $T'(\beta)$ as the iterative composite over the generators. We are then left to check that the relations of $\mathbb{B}$ are preserved (which they are). This then allows us to define $T' : \mathbb{B} \longrightarrow \mathcal{V}$ to be a unique, well defined strict monoidal functor which allows the diagram to commute. □

**Proposition 7.6.5.** Let $\mathcal{V}$ be a strict monoidal category, and suppose we have a functor $T : \mathcal{A} \longrightarrow \mathcal{V}$ with associated Yang-Baxter operator $y$. Let $z$ be the Yang-Baxter operator on $i_{\mathcal{A}} : \mathcal{A} \longrightarrow \mathbb{B} \int \mathcal{A}$. Then there exists a unique strict monoidal functor $T' : \mathbb{B} \int \mathcal{A} \longrightarrow \mathcal{V}$ such that the diagram

$$\begin{tikzcd}
\mathcal{A} \arrow[r, "i_{\mathcal{A}}"] \arrow[dr, "T"'] & \mathbb{B} \int \mathcal{A} \arrow[d, dashed, "T'"] \\
& \mathcal{V}
\end{tikzcd}$$

commutes and that $T'(y') = y$.

*Proof.* Recall that $\{\mathcal{A}, \mathcal{V}\}$ is a strict monoidal category if $\mathcal{V}$ is. Consider again the one point category $\mathbf{1}$ and construct functors $F_S : \mathbf{1} \longrightarrow \{\mathcal{A}, \mathcal{V}\}$ and $j : \mathbf{1} \longrightarrow \mathbb{B}$ where $F_T(\bullet) = T : \mathcal{A} \longrightarrow \mathcal{V}$ and $i(\bullet) = 1$. Then by the previous work, there exists a map $T^{\#} : \mathbb{B} \longrightarrow \{\mathcal{A}, \mathcal{V}\}$ such that the diagram below commutes.

$$\begin{tikzcd}
\mathbf{1} \arrow[r, "i"] \arrow[dr, "F_T"'] & \mathbb{B} \arrow[d, dashed, "T^{\#}"] \\
& \{\mathcal{A}, \mathcal{V}\}
\end{tikzcd}$$

Now construct the maps $\{F_S\} : \{*\} \longrightarrow \text{Hom}(\mathbf{1}, \{\mathcal{A}, \mathcal{V}\})$ and $\{S\} : \{*\} \longrightarrow \text{Hom}(\mathcal{A}, \mathcal{V})$ where $\{F_S\}(*) = F_S$ and $\{S\}(*) = S$. Consider the pullback squares below.

$$\begin{tikzcd}
P \arrow[r] \arrow[d] & \text{Hom}(\mathbb{B}, \{\mathcal{A}, \mathcal{V}\}) \arrow[d, "(-)\circ i"] \\
\{*\} \arrow[r, "\{F_S\}"'] & \text{Hom}(\mathbf{1}, \{\mathcal{A}, \mathcal{V}\})
\end{tikzcd}
\qquad
\begin{tikzcd}
(K, T) \arrow[r, mapsto] \arrow[d, mapsto] & T : \mathbb{B} \longrightarrow \{\mathcal{A}, \mathcal{V}\} \arrow[d, mapsto] \\
* \arrow[r, mapsto] & F_S = T \circ i
\end{tikzcd}$$

$$\begin{tikzcd}
Q \arrow[r] \arrow[d] & \text{Hom}(\mathbb{B} \int \mathcal{A}, \mathcal{V}) \arrow[d, "(-)\circ i_A"] \\
\{*\} \arrow[r] & \text{Hom}(\mathcal{A}, \mathcal{V})
\end{tikzcd}
\qquad
\begin{tikzcd}
(K', T') \arrow[r, mapsto] \arrow[d, mapsto] & T' : \mathbb{B} \longrightarrow \{\mathcal{A}, \mathcal{V}\} \arrow[d, mapsto] \\
* \arrow[r, mapsto, "\{S\}"'] & S = T' \circ i_A
\end{tikzcd}$$

First, $P$ corresponds to the set of functors $T : \mathbb{B} \longrightarrow \{\mathcal{A}, \mathcal{V}\}$ such that precomposition with $i$ is equal to $F$. Meanwhile, the set $Q$ consists of functors $T' : \mathbb{B} \int \mathcal{A} \longrightarrow \mathcal{V}$ where precomposition with $i_{\mathcal{A}}$ is equal to $S$. However, these sets are in bijection due to the adjoint relation we have. In other words, the diagrams

$$\begin{tikzcd}
\mathbf{1} \arrow[r, "i"] \arrow[dr, "F_S"'] & \mathbb{B} \arrow[d, dashed, "T"] \\
& \{\mathcal{A}, \mathcal{V}\}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathcal{A} \arrow[r, "i_{\mathcal{A}}"] \arrow[dr, "S"'] & \mathbb{B} \int \mathcal{A} \arrow[d, dashed, "T'"] \\
& \mathcal{V}
\end{tikzcd}$$

are in bijection. Hence we see that $T^{\#}$ corresponds uniquely with a functor $T'$ such that the diagram

$$\begin{tikzcd}
\mathcal{A} \arrow[r, "i_{\mathcal{A}}"] \arrow[dr, "T"'] & \mathbb{B} \int \mathcal{A} \arrow[d, dashed, "T'"] \\
& \mathcal{V}
\end{tikzcd}$$

commutes and preserves the Yang-Baxter operators as desired. □

**Theorem 7.6.6.** Let $\mathcal{V}$ be an $B$-category and suppose we have a functor $F : \mathcal{A} \longrightarrow \mathcal{V}$. Then there is an equivalence of categories

$$\mathbb{B}Fun(\mathbb{B} \int \mathcal{A}, \mathcal{V}) \simeq Fun(\mathcal{A}, \mathcal{V}).$$

given by precomposition of each $F : \mathbb{B} \int \mathcal{A} \longrightarrow \mathcal{V}$ with $i_{\mathcal{A}} : \mathcal{A} \longrightarrow \mathbb{B} \int \mathcal{A}$.

*Proof.* We follow the same argument as Joyal and Street. By the previous lemma, every $SB$-monoidal category is strongly equivalent to a strict $SB$-monoidal category $\mathcal{V}'$ via a pair of functors $E : \mathcal{V} \longrightarrow \mathcal{V}'$ and $E' : \mathcal{V}' \longrightarrow \mathcal{V}$. Hence observe that if we have an equivalence of categories $(-) \circ i_{\mathcal{A}} : \mathbb{B}\,\mathrm{Fun}(\mathbb{B} \int \mathcal{A}, \mathcal{V}') \longrightarrow \mathrm{Fun}(\mathcal{A}, \mathcal{V}')$, then the diagram below commutes

$$
\begin{tikzcd}
\mathbb{B}\mathrm{Fun}(\mathbb{B}\int\mathcal{A}, \mathcal{V}) \arrow[r, dashed, "(-)\circ i_{\mathcal{A}}"] \arrow[d, "F\circ(-)"'] & \mathrm{Fun}(\mathcal{A}, \mathcal{V}) \\
\mathbb{B}\mathrm{Fun}(\mathbb{B}\int\mathcal{A}, \mathcal{V}') \arrow[r, "(-)\circ i_{\mathcal{A}}"'] & \mathrm{Fun}(\mathcal{A}, \mathcal{V}') \arrow[u, "E'\circ(-)"']
\end{tikzcd}
$$

and the top dashed arrow is an equivalence as well. So it suffices to prove this for the strict case. Now, the proposed functor $F$ behaves as

$$F(S : \mathbb{B} \int \mathcal{A} \longrightarrow \mathcal{V}) = S \circ i_{\mathcal{A}} : \mathcal{A} \longrightarrow \mathcal{V}.$$

We must demonstrate that this is fully faithful and essentially surjective.

**Fully faithful.** Let $F, G : \mathbb{B} \int \mathcal{A} \longrightarrow \mathcal{V}$ be strong $SB$-monoidal functors. Then define the function

$$\varphi : \mathrm{Hom}_{\mathbb{B}\mathrm{Fun}(\mathbb{B}\int\mathcal{A}, \mathcal{V})}(F, G) \longrightarrow \mathrm{Hom}_{\mathrm{Fun}(\mathcal{A},\mathcal{V})}(F \circ i_{\mathcal{A}}, G \circ_{i_{\mathcal{A}}}).$$

where, given a natural transformation $\eta : F \longrightarrow G$, we have that $\varphi(\eta) : F \circ i_{\mathcal{A}} \longrightarrow G \circ i_{\mathcal{A}}$ is a natural transformation defined as

$$\varphi(\eta)_A = \eta_{[A]}.$$

We show that this is injective. Suppose $\varphi(\eta) = \varphi(\eta')$ for two natural transformations $\eta, \eta' : F \longrightarrow G$ with $F, G \in \mathbb{B}\mathrm{Fun}(\mathbb{B} \int \mathcal{A}, \mathcal{V})$. The fact that $\varphi(\eta) = \varphi(\eta')$ implies that

$$\eta_{[A]} = \eta_{[A']}.$$

As these are natural transformations between monoidal functors, we have that the diagram below commutes.

$$
\begin{tikzcd}
F([A_1]) \otimes \cdots \otimes F([A_n]) \arrow[r, "\eta_{[A_1]}\otimes\cdots\otimes\eta_{[A_n]}"] \arrow[d, "P_1"', "\cong"] & G([A_1]) \otimes \cdots \otimes G([A_n]) \arrow[d, "\cong"', "P_2"] \\
F([A_1, \dots, A_n]) \arrow[r, "\eta_{[A_1,\dots,A_n]}"'] & G([A_1, \dots, A_n])
\end{tikzcd}
$$

The morphisms $P_1$ and $P_2$ are the isomorphisms built inductively from

$$F_2 : F([A]) \otimes F([B] \xrightarrow{\sim} F([A, B])$$

which comes equipped with the data of a strong monoidal functor [see Mac Lane, p. 256]. Moreover, the diagram commutes by Mac Lane's coherence theorem.
The above diagram similarly holds with $\eta$ replaced as $\eta'$, since $\eta'$ is also a natural transformation of monoidal functors. Hence what we see is that

$$\begin{aligned}
\eta_{[A_1,\dots,A_n]} \circ P_1 &= P_2 \circ \eta_{[A_1]} \otimes \cdots \otimes \eta_{[A_n]}\\
&= P_2 \circ \eta'_{[A_1]} \otimes \cdots \otimes \eta'_{[A_n]}\\
&= \eta'_{[A_1,\dots,A_n]} \circ P_1.
\end{aligned}$$

As $P_1$ is an isomorphism, we have that $\eta_{[A_1,\dots,A_n]} = \eta'_{[A_1,\dots,A_n]}$, so that $\varphi(\eta) = \varphi(\eta')$ implies that $\eta = \eta'$. Hence the functor is faithful. The functor is clearly full, since by the above process we can always take a natural transformation $\eta : F \circ i_{\mathcal{A}} \longrightarrow G \circ i_{\mathcal{A}}$ and build it into a natural transformation $\eta : F \longrightarrow G$.

**Essentially Surjective.** Consider a functor $F : \mathcal{A} \longrightarrow \mathcal{V}$. By Proposition 7.6.5, we know there exists a unique $S : \mathbb{B} \int \mathcal{A} \longrightarrow \mathcal{V}$ such that $S \circ i_{\mathcal{A}} = F$. Hence we have essential surjectivity; in fact, we have a stronger version in the strict case.

□

# 7.7 Monoids, Groups, in Symmetric Monoidal Categories

Recall from section ? that we were able to construct monoid and groups which were internal to some category $\mathcal{C}$. The philosophy behind the construction is one we've seen before: we of course think of monoids and groups by their elements, but we resist the temptation and instead present an object-free, diagrammatic set of axioms for monoids and rings. We utilized the cartesian product in the category $\mathcal{C}$ to demonstrate this. However, we now know that the cartesian product in any category is a small example of a category with a symmetric monoidal structure. Hence we revisit the concepts of a monoid and group, and expand their generality by demonstrating that they can be defined in a symmetric monoidal category.

**Definition 7.7.1.** Let $(\mathcal{M}, \otimes, I, \alpha, \rho, \lambda)$ be a monoidal category and let $M$ be an object of $\mathcal{M}$. We say $M$ is if there exist maps

$$\mu : M \otimes M \longrightarrow M$$
$$\eta : I \longrightarrow M$$

referred to as the multiplication and identity maps, such that the diagrams below commute.

$$\begin{tikzcd}
M \otimes (M \otimes M) \arrow[r, "\alpha"] \arrow[d, "1\otimes\mu"'] & (M \otimes M) \otimes M \arrow[r, "\mu\otimes 1_M"] & M \otimes M \arrow[d, "\mu"] \\
M \otimes M \arrow[rr, "\mu"'] & & M
\end{tikzcd}$$

$$\begin{tikzcd}
I \otimes M \arrow[r, "\eta\otimes 1_M"] \arrow[dr, "\lambda_M"'] & M \otimes M \arrow[d, "\mu"] & M \otimes I \arrow[l, "1_M\otimes\eta"'] \arrow[dl, "\rho_M"] \\
& M &
\end{tikzcd}$$

---

**Example 7.7.2.** One of the most useful examples of this concept arises from the notion of an algebra $A$ over some field $k$, where $A$ is a vector space over the field $k$.

---

## 7.8 Enriched Categories

When we originally defined categories, we sought a degree of large generality that was able to capture a huge amount of mathematical phenomenon. However, this was not out a mere desired for generality; as Mac Lane puts it, "good general theory does not search for the maximum generality, but for the right generality" (108). But it does turn out that in defining categories so widely we lose some of their internal structure; for example, in many categories, every homset might have a underlying abelian group structure. These are called **preadditive categories** and are extremely useful, in that they give us a first step towards a general framework (but not to general) that allows one to do homological algebra in.

Now if we've lost some original framework, how do we recover it? First, recall that in categories, objects are basically dummies. It doesn't matter how I denote my objects in my category $\mathcal{C}$; you and are I talking about the same category if our morphisms act the same exact way. For example, the categories

$$
\begin{tikzcd}
1 \arrow[r] \arrow[loop above] & 2 \arrow[r] \arrow[loop above] & 3 \arrow[r] \arrow[loop above] & \cdots \arrow[r] & n \arrow[loop above]
\end{tikzcd}
$$

and

$$
\begin{tikzcd}
\text{All} \arrow[r] \arrow[loop above] & \text{Politicians} \arrow[r] \arrow[loop above] & \text{Are} \arrow[r] \arrow[loop above] & \cdots \arrow[r] & \text{Corrupt} \arrow[loop above]
\end{tikzcd}
$$

where the above objects are $n$ words describing how politicians suck, are the same preorders. Thus, because categorical structure is primarily found within the morphisms, i.e. the homsets, we only need to fix these to take back our original structure.

**Definition 7.8.1.** Let $(\mathcal{V}, \otimes, I)$ be a monoidal category. A small category $\mathcal{C}$ is a $\mathcal{V}$-category or an **enriched category** over $\mathcal{V}$ if

1. For each $A, B \in \mathcal{C}$, we have that $\text{Hom}_{\mathcal{C}}(A, B) \in \mathcal{V}$
2. There exists a "composition" operator

$$\circ_{A,B,C} : \text{Hom}_{\mathcal{C}}(A, B) \times \text{Hom}_{\mathcal{C}}(B, C) \longrightarrow \text{Hom}_{\mathcal{C}}(A, C)$$

3. For each object $A \in \mathcal{C}$, we have a "identity object"

$$i_A : I \longrightarrow \text{Hom}_{\mathcal{C}}(A, A)$$

such that our composition operator is associative:

$$
\begin{tikzcd}
\mathrm{Hom}(A,B)\otimes(\mathrm{Hom}(B,C)\otimes\mathrm{Hom}(C,D)) \arrow[rr, "\alpha"] \arrow[d, "1\otimes\circ_{B,C,D}"'] & & (\mathrm{Hom}(A,B)\otimes\mathrm{Hom}(B,C))\otimes\mathrm{Hom}(C,D) \arrow[d, "\circ_{A,B,C}\otimes 1"] \\
\mathrm{Hom}(A,B)\otimes\mathrm{Hom}(B,D) \arrow[r, "\circ_{A,B,D}"'] & \mathrm{Hom}(A,D) & \mathrm{Hom}(A,C)\otimes\mathrm{Hom}(C,D) \arrow[l, "\circ_{B,C,D}"]
\end{tikzcd}
$$

and such that our unital elements in each homset behave morally like an identity element should:

$$
\begin{tikzcd}
\mathrm{Hom}(B,B)\otimes\mathrm{Hom}(A,B) \arrow[r, "\circ_{A,B,B}"] & \mathrm{Hom}(A,B) & \mathrm{Hom}(A,B)\otimes\mathrm{Hom}(A,A) \arrow[l, "\circ_{A,A,B}"'] \\
I\otimes\mathrm{Hom}(A,B) \arrow[u, "i_B\otimes 1"] \arrow[ur, "\lambda"'] & & \mathrm{Hom}(A,B)\otimes I \arrow[u, "1\otimes i_A"'] \arrow[ul, "\rho"]
\end{tikzcd}
$$

---

**Example 7.8.2.** The following is a classic example due to F.W. Lawvere. A **Lawvere metric space** is a set $X$ equipped with a distance function $d : X \times X \longrightarrow \mathbb{R}$ such that

1. $d(x, x) = 0$ for all $x \in X$
2. $d(x, z) \le d(x, y) + d(y, z)$ for all $x, y, z \in X$.

It turns out that, we may equivalently define such a space as a category enriched over $([0, \infty), +, 0)$.

Recall that $([0, \infty), +, 0)$ where $+$ is addition forms a symmetric monoidal category. Here we treat $[0, \infty]$ as a poset where for a pair of objects $a, b$ there exists exactly one morphism

$$a \longrightarrow b \text{ iff } b \le a.$$

Now what does it look like for a category $\mathcal{C}$ to be $[0, \infty]$-category? It means that for any pair of objects $A, B$, we have that $\mathrm{Hom}_{\mathcal{C}}(A, B) \in [0, \infty)$. If we denote $d(A, B) = \mathrm{Hom}_{\mathcal{C}}(A, B)$, this then implies that we have a function

$$d : \mathrm{Ob}(\mathcal{C}) \times \mathrm{Ob}(\mathcal{C}) \longrightarrow [0, \infty].$$

Enriched categories also grant us a composition morphism

$$\mathrm{Hom}_{\mathcal{C}}(A, B) \times \mathrm{Hom}_{\mathcal{C}}(B, C) \longrightarrow \mathrm{Hom}_{\mathcal{C}}(A, C)$$

for all objects $A, B, C$. But in $[0, \infty)$, morphisms are just size relations, so what this really means is that

$$d(A, C) \le d(A, B) + d(B, C)$$

for all $A, B, C \in \mathcal{C}$ Finally, we see the identity criterion states that for each object $A$, we have a morphism $i_A : 0 \longrightarrow \mathrm{Hom}_{\mathcal{C}}(A, A)$ which translates to

$$d(A, A) \le 0 \implies d(A, A) = 0$$

since $d(A, A) \in [0, \infty]$. This should feel very familiar; what we've just come up with is nearly a metric space structure on the objects of our category! We are only missing the symmetry relation. For that, this special construction is known as a **Lawvere metric space**.

---

---

**Example 7.8.3.** Recall that a (strict) 2-category is a category $\mathcal{C}$ such that, in addition to the morphisms $f : A \longrightarrow B$ between objects $A, B \in \mathcal{C}$, there exists 2-morphisms $\alpha : f \longrightarrow g$ between parallel morphisms $f, g : A \longrightarrow B$.

$$\begin{tikzcd} A \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] \arrow[r, phantom, "\Downarrow\alpha"] & B \end{tikzcd}$$

These two morphisms have access to two different forms of composition. On one hand, there is "vertical" composition

$$\begin{tikzcd} A \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] \arrow[r, phantom, "\Downarrow\alpha"] & B \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} A \arrow[r, bend left, "g"] \arrow[r, bend right, "h"'] \arrow[r, phantom, "\Downarrow\beta"] & B \end{tikzcd} \quad \Longrightarrow \quad \begin{tikzcd} A \arrow[r, bend left, "f"] \arrow[r, bend right, "h"'] \arrow[r, phantom, "\Downarrow\beta\bullet\alpha"] & B \end{tikzcd}$$

while on the other, there is "horizontal" composition.

$$\begin{tikzcd} A \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] \arrow[r, phantom, "\Downarrow\gamma"] & B \arrow[r, bend left, "h"] \arrow[r, bend right, "k"'] \arrow[r, phantom, "\Downarrow\delta"] & C \end{tikzcd} \quad \Longrightarrow \quad \begin{tikzcd} A \arrow[r, bend left, "h\circ f"] \arrow[r, bend right, "k\circ g"'] \arrow[r, phantom, "\Downarrow\delta\circ\gamma"] & B \end{tikzcd}$$

Moreover, we require that the interchange law be satisfied and that the morphisms form a category under the vertical composition given by $\circ$. However, we can rephrase this as saying a category $\mathcal{C}$ is a 2-category if

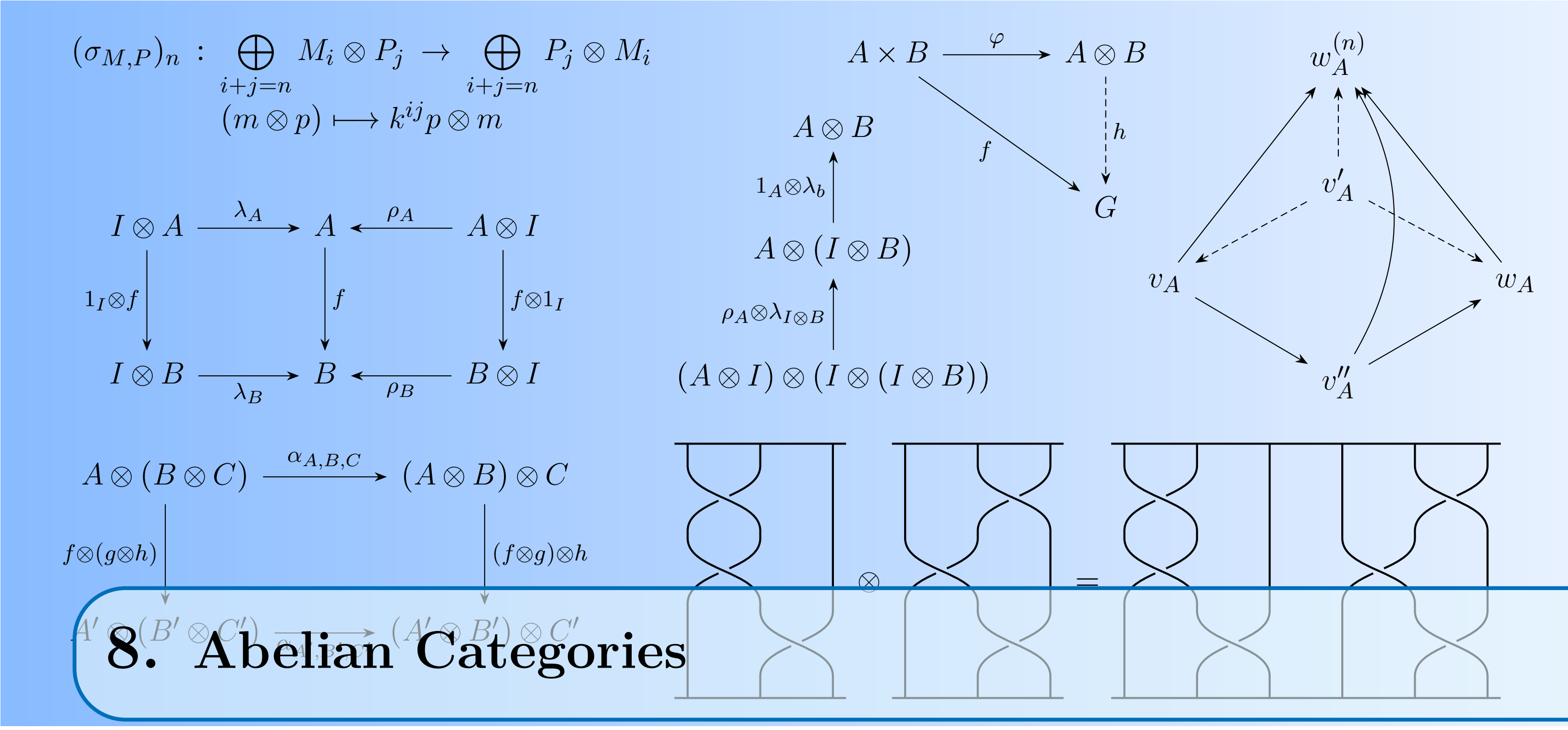
1. For each $A, B \in \mathcal{C}$ we have that $(\mathrm{Hom}_{\mathcal{C}}(A, B), \circ)$ is a category
2. There exist a composition operator $\circ : \mathrm{Hom}(A, B) \times \mathrm{Hom}(B, C) \longrightarrow \mathrm{Hom}(A, C)$

$$\begin{tikzcd} A \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] \arrow[r, phantom, "\Downarrow\gamma"] & B \arrow[r, bend left, "h"] \arrow[r, bend right, "k"'] \arrow[r, phantom, "\Downarrow\delta"] & C \end{tikzcd} \quad \Longrightarrow \quad \begin{tikzcd} A \arrow[r, bend left, "h\circ f"] \arrow[r, bend right, "k\circ g"'] \arrow[r, phantom, "\Downarrow\delta\circ\gamma"] & B \end{tikzcd}$$

3. For each object $A$, we have a functor $i_A : 1 \longrightarrow \mathrm{Hom}(A, A)$, where 1 is the one object category with one morphism that is sent to $1_A$.

Above, (3) is stupidly simple; but the reason we're framing it this way is to demonstrate that a strict 2-category $\mathcal{C}$ is the same thing as a category $\mathcal{C}$ enriched over the monoidal category $(\mathbf{Cat}, \times, 1)$; the category of small categories whose monoidal product is the cartesian product and whose identity is the one-object-one-morphism category 1.

---



# 8. Abelian Categories

Abelian categories are generalizations of the structure which can be found in the category of abelian groups **Ab**. This may be obvious from the name; what is nontrivial, however, is how to preserve the nice structure of the category without specific reference to the elements themselves. It turns out this is possible, but is generally not the way we think about **Ab**. This is the aim of this chapter.

## 8.1 Preadditive Categories

Consider two abelian groups $(G, +)$ and $(H, \cdot)$ of **Ab**. Recall from group theory that we can turn the set $\text{Hom}(G, H)$ into an abelian group $(\text{Hom}(G, H), *)$ as follows. Given $\varphi, \psi : G \longrightarrow H$, we can create another group homomorphism $\varphi * \psi : G \longrightarrow H$ where

$$(\varphi * \psi)(g) = \varphi(g) \cdot \psi(g).$$

Observe that this is in fact a group homomorphism: if $g, g' \in G$, then

$$\begin{aligned}
(\varphi * \psi)(g + g') &= \varphi(g + g') \cdot \psi(g + g') \\
&= \varphi(g) \cdot \varphi(g') \cdot \psi(g) \cdot \psi(g') \\
&= \varphi(g) \cdot \psi(g) \cdot \varphi(g') \cdot \psi(g') \\
&= (\varphi * \psi)(g) \cdot (\varphi * \psi)(g').
\end{aligned}$$

In the third step we utilized the fact that $(H, \cdot)$ is abelian. Thus $(\text{Hom}(G, H), *)$ is not necessarily a group unless $H$ is an abelian group. Therefore, this construction doesn't extend to **Grp**.

At this point, your category-theory-voice in your head is probably asking:

If $H$ is an abelian group, can we create a functor $F_H : \mathbf{Ab} \longrightarrow \mathbf{Ab}$ where $G \mapsto \text{Hom}(G, H)$?

The answer is yes; the functor is actually contravariant, for suppose we have a group homomorphism

$$\varphi : G \longrightarrow G'.$$

Then define the function

$$F_H(\varphi) : \text{Hom}(G', H) \longrightarrow \text{Hom}(G, H)$$

where

$$F_H(\varphi)(\psi : G' \longrightarrow H) = \psi \circ \varphi : G \longrightarrow H.$$

To verify functoriality, we have to check that this function is actually a group homomorphism. Towards that goal, consider $\psi, \sigma : G \longrightarrow H$. Then observe that for any $g \in G$,

$$\begin{aligned} F_H(\varphi)(\psi + \sigma)(g) &= \varphi(\psi(g) + \sigma(g)) \\ &= \varphi(\psi(g)) + \varphi(\sigma(g)) \\ &= F_H(\varphi)(\psi)(g) + F_H(\varphi)(\psi)(g) \end{aligned}$$

which verifies that $F_H(\varphi)$ is a group homomorphism. Therefore, we see that $F_H : \mathbf{Ab} \longrightarrow \mathbf{Ab}$ is in fact a functor.

Now your category-theory-voice should be asking:

If $G$ is an abelian group, can we *also* create a functor $F^G : \mathbf{Ab} \longrightarrow \mathbf{Ab}$ where $H \mapsto \text{Hom}(G, H)$?

One can easily show that the answer is yes. In this direction, the functor is covariant. That is, for $\psi : H \longrightarrow H'$, we have that

$$F^G(\psi) : \text{Hom}(G, H) \longrightarrow \text{Hom}(G, H')$$

where

$$F^G(\psi)(\varphi : G \longrightarrow H) = \psi \circ \varphi : G \longrightarrow H'.$$

Note that for our functors, we have that

$$F_H(G) = F^G(H).$$

This is *bifunctor-ish*. Therefore, our category theory voice is now asking:

Do we have a bifunctor $F : \mathbf{Ab} \times \mathbf{Ab} \longrightarrow \mathbf{Ab}$ on our hands, where $F(G, H) = \text{Hom}(G, H)$?

To see if this answer is true, we ought to be able to show that, given $\varphi : G' \longrightarrow G$ and $\psi : H \longrightarrow H'$, the diagram

$$\begin{tikzcd} F_H(G) \arrow[r, "F_H(\varphi)"] \arrow[d, "F^G(\varphi)"'] & F_H(G) \arrow[d, "F^{G'}(\psi)"] \\ F^{G'}(H) \arrow[r, "F_H(\psi)"'] & F^{G'}(H) \end{tikzcd} = \begin{tikzcd} \text{Hom}(G, H) \arrow[r, "(-)\circ\varphi"] \arrow[d, "\psi\circ(-)"'] & \text{Hom}(G', H) \arrow[d, "\psi\circ(-)"] \\ \text{Hom}(G, H') \arrow[r, "(-)\circ\varphi"'] & \text{Hom}(G', H') \end{tikzcd}$$

is commutative. The above diagram is in fact commutative since function composition is associative! That is, given $\sigma : G \longrightarrow H$, observe that going right and then down gives

$$\psi \circ (\sigma \circ \varphi)$$

while going down and then right gives

$$(\psi \circ \sigma) \circ \varphi.$$

Hence we have commutativity of the above diagram, and we therefore have a true bifunctor $F : \mathbf{Ab} \times \mathbf{Ab} \longrightarrow \mathbf{Ab}$ where

$$F(G, H) = \text{Hom}(G, H).$$

What this really shows is that $\text{Hom}(-, -)$ is a functor; specifically, a bifunctor. So while we typically think of $\text{Hom}(G, H)$ as a set, it had hidden functorial properties. Thus what makes **Ab** special is that plugging in abelian groups outputs an abelian group, and this is not the case with other constructions (e.g. **Grp**).

Let us now consider a new observation of **Ab**. For any triple of abelian groups

$$(G, \star), (H, +), (K, \cdot)$$

we can create abelian groups

$$\begin{aligned} &\Big(\text{Hom}(G, H), +'\Big) && (\varphi_1 +' \varphi_2)(g) = \varphi_1(g) + \varphi_2(g) \\ &\Big(\text{Hom}(H, K), \cdot'\Big) && (\psi_1 \cdot' \psi_2)(h) = \psi_1(h) \cdot \psi_2(h) \\ &\Big(\text{Hom}(G, K), *\Big) && (\sigma_1 * \sigma_2)(g) = \sigma_1(g) \cdot \sigma_2(g) \end{aligned}$$

where $\varphi_i \in \text{Hom}(G, H), \psi_i \in \text{Hom}(H, K)$ and $\sigma_i \in \text{Hom}(G, K)$ for $i = 1, 2$. Now since these are abelian groups in **Ab**, there is a composition operator

$$\circ : \text{Hom}(G, H) \times \text{Hom}(H, K) \longrightarrow \text{Hom}(G, K)$$

where $\circ(\varphi : G \longrightarrow H, \psi : H \longrightarrow K) \mapsto \psi \circ \varphi : G \longrightarrow K$. However, we now run into a problem where our operators might not play nicely with each other. Specifically, is it true that

$$\psi \circ (\varphi_1 +' \varphi_2) = (\psi \circ \varphi_1) * (\psi \circ \varphi_2)$$

or

$$(\psi_1 \cdot' \psi_2) \circ \varphi = (\psi_1 \circ \varphi) * (\psi_2 \circ \varphi)?$$

For the first case, the answer is yes. Observe that

$$\begin{aligned}
\psi \circ (\varphi_1 +' \varphi_2)(g) &= \psi(\varphi_1(g) + \varphi_2(g))\\
&= \psi(\varphi_1(g) + \varphi_2(g))\\
&= \psi(\varphi_1)(g) \cdot \psi(\varphi_2)(g)\\
&= \big((\psi \circ \varphi_1) * (\psi \circ \varphi_2)\big)(g).
\end{aligned}$$

The reason we have linearity here is because of the way we defined the **group operations** on the homsets. The definition of these operations is intuitively correct, but we get accidentally get an extra bonus of obtaining linearity so that we don't have to worry about the above equations not holding.

In order to mimic this behavior, we abstract this into a category to define a **Ab**-category.

**Definition 8.1.1.** An **Ab**-category or **Preadditive Category** is a category $\mathcal{C}$ such that, for each pair of objects $A, B$, there exists an abelian group operation $+$ on the set $\text{Hom}(A, B)$ such that

$$\begin{aligned}
&\circ : \text{Hom}(A, B) \times \text{Hom}(B, C) \longrightarrow \text{Hom}(A, C)\\
&(f, g) \mapsto g \circ f
\end{aligned}$$

is bilinear. What we mean by bilinear is that, given morphisms $f, g : A \longrightarrow B$ and $h, k : B \longrightarrow C$, we have that

$$\begin{aligned}
(h + k) \circ f &= h \circ f + k \circ f\\
h \circ (g + f) &= h \circ g + h \circ f.
\end{aligned}$$

Note that since we demand that $\text{Hom}_{\mathcal{C}}(A, B)$ always be a group, we see that any category such that $\text{Hom}_{\mathcal{C}}(A, B) = \varnothing$ can never be an abelian group. A group always requires the existence of an identity; a demand that an empty set can never meet. Therefore, as an example, any discrete category cannot be a preadditive category because all of the nontrivial homsets are empty.

As we demonstrated building up to this definition, **Ab** is a trivial example of a preadditive category. A less trivial example is $\mathbf{Vect}_K$ where $K$ is a field, but this is nearly automatic since this takes advantage of the fact that vector spaces have their own hidden abelian group structure.

**Example 8.1.2.** Suppose $\mathcal{C}$ is a one object category $R$ which is also preadditive. Then this means that we have two binary operations $+$ and $\circ$ on the abelian group $\text{Hom}_{\mathcal{C}}(R, R)$ such that

$$(h + k) \circ f = h \circ f + k \circ f$$
$$h \circ (g + f) = h \circ g + h \circ f.$$

However, this is simply a ring! The addition is the ring addition, while the ring multiplication is given by composition. Conversely, a ring regarded as the homset of a one object category can be defined to be an abelian category. This is because when regarding a group as a one object category, the group operation becomes the composition operation. Thus adding the extra axiom of an addition bilinear operation grants us that the category is preadditive.

**Example 8.1.3.** Let $\mathcal{C}$ be a preadditive category. Then $\mathcal{C}^{\text{op}}$ is also a preadditive category. To demonstrate this, we know that every pair of objects $A, B \in \mathcal{C}$ gives rise to a group $(\text{Hom}_{\mathcal{C}}(A, B), +)$ for some operation $+$. This allows us to place a group structure $+'$ on $\text{Hom}_{\mathcal{C}^{\text{op}}}(B, A)$ where for two $f^{\text{op}}, g^{\text{op}} : B \longrightarrow A$ in $\mathcal{C}^{\text{op}}$,

$$f^{\text{op}} +' g^{\text{op}} = (f + g)^{\text{op}}.$$

That is, we rely on the preexisting group operation $+$ from $\text{Hom}_{\mathcal{C}}(A, B)$. Given that the composition operator of $\mathcal{C}^{\text{op}}$ is $\circ^{\text{op}}$, we can check that this satisfies the bilinearity conditions of $\circ^{\text{op}}$. Suppose $h^{\text{op}}, k^{\text{op}} : B \longrightarrow A$ are two morphisms in $\text{Hom}(B, A)$ which are composable with some $f^{\text{op}}$. Then

$$\begin{aligned}(h^{\text{op}} +' k^{\text{op}}) \circ^{\text{op}} f^{\text{op}} = (h + k)^{\text{op}} \circ^{\text{op}} f^{\text{op}} &= f \circ (h + k)\\ &= f \circ h + f \circ k\\ &= h^{\text{op}} \circ^{\text{op}} f^{\text{op}} +' k^{\text{op}} \circ^{\text{op}} f^{\text{op}}.\end{aligned}$$

The other direction can be verified dually, so that the the group operation $+'$ distributes bilinearly over $\circ^{\text{op}}$. Therefore, $\mathcal{C}^{\text{op}}$ is a preadditive category.

**Example 8.1.4.** If $\mathcal{C}$ is preadditive, then the functor category $\mathcal{C}^J$ is preadditive. To demonstrate this, consider the hom-set $\text{Hom}_{\mathcal{C}^J}(F, G)$ between two functors $F, G : J \longrightarrow \mathcal{C}$. Now consider two natural transformations $\eta, \varepsilon \in \text{Hom}_{\mathcal{C}^J}(F, G)$. Then for each $f \in \text{Hom}_{\mathcal{C}}(A, B)$, the familiar diagram commutes.

$$\begin{tikzcd} A \arrow[d, "f"] & F(A) \arrow[r, shift left, "\eta_A"] \arrow[r, shift right, "\varepsilon_A"'] \arrow[d, "F(f)"'] & G(A) \arrow[d, "G(A)"] \\ B & F(B) \arrow[r, shift left, "\eta_B"] \arrow[r, shift right, "\varepsilon_B"'] & G(B) \end{tikzcd}$$

This diagram tells us that $G(f) \circ \eta_A = \eta_B \circ F(f)$ and that $G(f) \circ \varepsilon_A = \varepsilon_B \circ F(f)$. However, since $\mathcal{C}$ is abelian, we can combine these morphisms and add both equations to get

$$G(f) \circ \eta_A + G(f) \circ \varepsilon_A = \eta_B \circ F(f) + \varepsilon_B \circ F(f) \implies G(f) \circ (\eta_A + \varepsilon_A) = (\eta_B + \varepsilon_B) \circ F(f).$$

Hence the diagram below

$$\begin{tikzcd} A \arrow[d, "f"] & F(A) \arrow[r, "\eta_A+\varepsilon_A"] \arrow[d, "F(f)"'] & G(A) \arrow[d, "G(A)"] \\ B & F(B) \arrow[r, "\eta_B+\varepsilon_B"'] & G(B) \end{tikzcd}$$

commutes. Therefore, using the group product of $(\mathrm{Hom}_{\mathcal{C}}(F(A), F(B)), +)$, we've derived a new natural transformation from $F$ to $G$ using $\eta$ and $\varepsilon$ in $\mathrm{Hom}_{\mathcal{C}^J}(F, G)$. This allows us to endow the homset $\mathrm{Hom}_{\mathcal{C}^J}(F, G)$ with the operation $+'$ defined so that for two $\eta, \varepsilon \in \mathrm{Hom}_{\mathcal{C}^J}(F, G)$, $\eta +' \varepsilon$ is the natural transformation where for each object $A$

$$(\eta +' \varepsilon)_A = \eta_A + \varepsilon_A$$

where $+$ is the group operation on $(\mathrm{Hom}_{\mathcal{C}}(F(A), G(A)), +)$. The fact that this distributes bilinearly over the composition operator is inherited from $\mathcal{C}$, and can easily be verified, so that $\mathcal{C}^J$ is preadditive.

---

---

**Example 8.1.5.** Let $\mathcal{C}$ be a category such that for every pair of objects $A, B$, the hom set $\mathrm{Hom}_{\mathcal{C}}(A, B)$ is nonempty. Then we can create the category $\mathrm{PreAdd}(\mathcal{C})$ where the objects are the same as $\mathcal{C}$, except each $\mathrm{Hom}_{\mathrm{PreAdd}(\mathcal{C})}(A, B)$ is now regarded as the free abelian group generated by the elements of $\mathrm{Hom}_{\mathcal{C}}(A, B)$. This results in a preadditive category if we force the composition operator $\circ'$ in $\mathrm{PreAdd}(\mathcal{C})$ to be bilinear. This forcing makes sense in our case since, if $\sum_{f \in \mathrm{Hom}_{\mathcal{C}}(A,B)} n_f f, \sum_{f \in \mathrm{Hom}_{\mathcal{C}}(A,B)} n'_f f$ are two arbitrary elements in $\mathrm{Hom}_{\mathrm{PreAdd}(\mathcal{C})}(A, B)$, then if $\sum_{k \in \mathrm{Hom}_{\mathcal{C}}(B,C)} m_k k \in \mathrm{Hom}_{\mathrm{PreAdd}(\mathcal{C})}(B, C)$ for some object $C$, where $n_f, n'_f, m_k$ are all nonzero for finitely many integers, then

$$\begin{aligned}
&\sum_{k\in\mathrm{Hom}_{\mathcal{C}}(B,C)} m_k k \circ' \left(\sum_{f\in\mathrm{Hom}_{\mathcal{C}}(A,B)} n_f f + \sum_{f\in\mathrm{Hom}_{\mathcal{C}}(A,B)} n'_f f\right)\\
&= \sum_{f\in\mathrm{Hom}_{\mathcal{C}}(A,B)} \sum_{k\in\mathrm{Hom}_{\mathcal{C}}(B,C)} n_f \cdot m_k(k\circ f) + \sum_{f\in\mathrm{Hom}_{\mathcal{C}}(A,B)} \sum_{k\in\mathrm{Hom}_{\mathcal{C}}(B,C)} n'_f \cdot m_k(k\circ f)
\end{aligned}$$

and the above last expression is in fact an element of $\mathrm{Hom}_{\mathrm{PreAdd}(\mathcal{C})}(A,C)$.

## 8.2 Additive Categories

Let $G$ and $H$ be abelian groups in **Ab**. A natural question to ask in any given category is if a binary product such at $G \times H$ exists in the category. In our case, the answer is yes; it is the **direct sum** $G \oplus H$. The direct sum satisfies the universal property

$$\begin{tikzcd} & K \arrow[dl, "\psi"'] \arrow[d, dashed, "u"] \arrow[dr, "\varphi"] & \\ G & G \oplus H \arrow[l, "\pi_G"] \arrow[r, "\pi_H"'] & H \end{tikzcd}$$

Here, $K$ is a third group, $\varphi$ and $\psi$ are arbitrary group homomorphisms, and $\pi_G, \pi_H$ are the natural projection morphisms. Interestingly, this object also satisfies the universal property

$$\begin{tikzcd} & K & \\ G \arrow[ur, "\psi"] \arrow[r, "i_G"'] & G \oplus H \arrow[u, dashed, "u"] & H \arrow[ul, "\varphi"'] \arrow[l, "i_H"] \end{tikzcd}$$

Here $i_G$ and $i_H$ are the natural injections, e.g. $i_G(g) = g \otimes e_H$. However, this implies that $G \oplus H$ is a coproduct! What this implies is that product and coproducts coincide in **Ab**. This is actually a pretty remarkable property because this isn't the case even in nice categories. For example, in **Set**, products and coproducts are definitely distinct.

Why is this the case?

**Proposition 8.2.1.** Let $\mathcal{C}$ be a preadditive category with a zero object $z$. Then for any objects $A, B \in \mathcal{C}$, the following are equivalent

$(i)$ $A \times B$ exists

$(ii)$ $A \amalg B$ exists

Moreover, there exists an isomorphism

$$\prod_{i \in \lambda} A_i \xrightarrow{\sim} \coprod_{i \in \lambda} A_i$$

for any objects $A_i \in \mathcal{C}$.

*Proof.* We only demonstrate one direction because the proof is self-dual.

Suppose $A \times B$ exists. Then then if $C$ is an object equipped with morphisms $f : C \longrightarrow A$ and $g : C \longrightarrow B$, the following diagram must hold.

$$\begin{tikzcd} & C \arrow[dl, "f"'] \arrow[d, dashed, "h"] \arrow[dr, "g"] & \\ A & A \times B \arrow[l, "\pi_A"] \arrow[r, "\pi_B"'] & B \end{tikzcd}$$

Equip $A$ with the morphisms $1_A : A \longrightarrow A$ and the unique zero morphism $\emptyset_A^B : A \longrightarrow B$. Then there exists a unique $i_A : A \longrightarrow A \times B$ such that the diagram commutes.

$$\begin{tikzcd} & A \arrow[dl, "1_A"'] \arrow[d, dashed, "i_A"] \arrow[dr, "\emptyset_A^B"] & \\ A & A \times B \arrow[l, "\pi_A"] \arrow[r, "\pi_B"'] & B \end{tikzcd}$$

Symmetrically, equip $B$ with the unique zero morphism $\emptyset_B^A : B \longrightarrow A$ and $1_B : B \longrightarrow B$. Then there exists a unique $i_B : B \longrightarrow A \times B$ such that the diagram commutes.

$$\begin{tikzcd} & B \arrow[dl, "\emptyset_B^A"'] \arrow[d, dashed, "i_B"] \arrow[dr, "1_B"] & \\ A & A \times B \arrow[l, "\pi_A"] \arrow[r, "\pi_B"'] & B \end{tikzcd}$$

Now we'll demonstrate that we have a coproduct structure on our hands. To do this, suppose we have an object $C$ equipped with morphisms $f : A \longrightarrow C$ and $g : B \longrightarrow C$. Then we can construct a morphism $h$ such that the following diagram commutes.

$$\begin{tikzcd} A \arrow[r, "i_A"] \arrow[dr, "f"'] & A \times B \arrow[d, dashed, "h"] & B \arrow[l, "i_B"'] \arrow[dl, "g"] \\ & C & \end{tikzcd}$$

Observe that $h = f \circ \pi_A + g \circ \pi_B$ suffices, where $+$ is the group operation on the abelian group $\mathrm{Hom}(A \times B, C)$. Observe that

$$\begin{aligned} h \circ i_A &= (f \circ \pi_A + g \circ \pi_B) \circ i_A \\ &= f \circ (\pi_A \circ i_A) + g \circ (\pi_B \circ i_A) \\ &= f. \end{aligned}$$

Similarly,

$$\begin{aligned} h \circ i_B &= (f \circ \pi_A + g \circ \pi_B) \circ 1_B \\ &= f \circ (\pi_A \circ 1_B) + g \circ (\pi_B \circ 1_B) \\ &= g. \end{aligned}$$

Hence the commutativity of the above diagram holds; therefore, we see that $A \times B$ is also a coproduct. Finally, recall that if two distinct objects satisfy the same universal property, they are necessarily isomorphic; therefore the existence of an isomorphism between the product and coproduct is immediate. ■

The above proof is not hard, but it's also not trivial. Moreover, there are three extremely important ingredients we utilized that demonstrate that the assumptions we've made so far are actually necessary and useful.

- This proof does not hold for a category without a zero object because there is not, in general, an obviously conceivable morphism to go from any two objects $A$ and $B$.
- Notice that calculating $h$ was only possible because we had an abelian group operation.
- Finally, notice that we utilized bilinearity of the composition operator in order to calculate $h \circ i_A$ and $h \circ i_B$ and thereby verify the universal property.

Therefore, all of our assumptions so far have been necessary and useful. And all of this now motivates the following definition.

**Definition 8.2.2.** Let $\mathcal{C}$ be an abelian category. A **biproduct** of two objects $A, B$ of $\mathcal{C}$ is an object $A \otimes B$ which is both a product and coproduct.
Equivalently, A biproduct is an object $A \oplus B$ equipped with morphisms

$$\pi_A : A \oplus B \longrightarrow A \qquad i_A : A \longrightarrow A \oplus B$$
$$\pi_B : A \oplus B \longrightarrow B \qquad i_B : B \longrightarrow A \oplus B$$

such that

1. $\pi_A \circ i_A = 1_A$
2. $\pi_B \circ i_B = 1_B$
3. $i_A \circ \pi_A + i_B \circ \pi_B = 1_{A \oplus B}$

**Definition 8.2.3.** An **Additive Category** is a preadditive category $\mathcal{C}$ such that finite biproducts exist.

**Definition 8.2.4.** Consider the category **Grp**.

# 8.3 Preabelian Categories

In **Ab**, kernels and cokernels exists for every group homomorphism.

First, recall their definitions.

**Definition 8.3.1.** Let $\varphi : G \longrightarrow H$ be a group homomorphism. Then a **kernel** is an equalizer of $\varphi : G \longrightarrow H$ and $0 : G \longrightarrow H$, where 0 maps everything to $e_H$, while a **cokernel** is a coequalizer of $\varphi : G \longrightarrow H$ and $0 : G \longrightarrow H$.

$$\begin{tikzcd} \mathrm{Ker}(\varphi) \arrow[r, "e"] & G \arrow[r, shift left, "\varphi"] \arrow[r, shift right, "0"'] & H \arrow[r, "c"] & \mathrm{Coker}(\varphi) \end{tikzcd}$$

In **Ab**, we set $\mathrm{Coker}(\varphi) \cong H/\operatorname{Im}(\varphi)$ while $\mathrm{Ker}(\varphi)$ is the natural normal subgroup of $G$.

Note that the necessary conditions for creating kernels and cokernels is (1) the existence of a zero object and (2) the existence of equalizers. If we have these ingredients, can we extend the concept of kernels and cokernels to additive categories? We can.

**Definition 8.3.2.** Let $\mathcal{C}$ be a category with a zero object as well as equalizers and coequalizers. Let $f : A \longrightarrow B$ be a morphism between two objects in $\mathcal{C}$. We define

- **kernel** to be the equalizer of $f$ and $\emptyset_A^B : A \longrightarrow B$, the zero morphism,
- **cokernel** of $f$ to be the coequalizer of $f$ and $\emptyset_A^B : A \longrightarrow B$.

In diagrams, we have that

$$\begin{tikzcd} \mathrm{Ker}(f) \arrow[r, "e"] & A \arrow[r, shift left, "f"] \arrow[r, shift right, "\emptyset"'] & B \arrow[r, "c"] \arrow[dr, "\psi"'] & \mathrm{Coker}(f) \arrow[d, dashed, "k"] \\ C \arrow[u, dashed, "h"] \arrow[ur, "\varphi"'] & & & D \end{tikzcd}$$

---

**Example 8.3.3.** In the category **Grp**, we certainly have a zero object $z = \{e\}$. Observe that for a given morphism $\varphi : G \longrightarrow H$, we can also form the equalizer of $\varphi$ by considering the pair $(\mathrm{Ker}(\varphi), e : \mathrm{Ker}(\varphi) \longrightarrow G)$ where $\mathrm{Ker}(\varphi) \subseteq G$ and $e$ being inclusion. For the same morphism, we can form the coequalizer be considering the pair $(\overline{N}, c : H \longrightarrow H/\overline{N})$ where

$$\overline{N} = \bigcap_{N \in \lambda} N$$

where $\lambda = \{H' \subseteq H \mid \operatorname{Im}(\varphi) \subseteq H' \text{ and } H' \trianglelefteq H\}$. It's a simple exercise to show that these satisfy the necessary universal properties.

However, it's important to observe the subtle difference between the behaviors of **Grp** and **Ab**. Because every subgroup of an abelian group is normal, we know that in the case of **Ab**, $\overline{N} = \operatorname{Im}(\varphi)$ So the coequalizer becomes

$$(\operatorname{Im}(\varphi), c : H \longrightarrow H/\operatorname{Im}(\varphi)).$$

---

It turns out that kernels and cokernels are extremely flexible in additive categories.

**Proposition 8.3.4.** Suppose $\mathcal{C}$ is an additive category. Then the following are equivalent.

$(i)$ $\mathcal{C}$ has equalizers and coequalizers.

$(ii)$ $\mathcal{C}$ has kernels and cokernels.

*Proof.* We only prove the statement for equalizers as the proof will be self-dual.

First note that $(i) \implies (ii)$ is immediate because a kernel is an equalizer with a morphism $\varphi$ and a zero morphism. To show $(ii) \implies (i)$, suppose that we have kernels for every morphism in $\mathcal{C}$. Then consider two morphisms $\varphi, \psi : G \longrightarrow H$. We can combine these two morphisms by our group operation on $\text{Hom}(G, H)$ and consider $\varphi - \psi$. Since we can take kernels, we take the kernel of this morphism.

$$\text{Ker}(\varphi) \xrightarrow{e} G \xrightarrow{\varphi-\psi} H$$

We now argue that this is the equalizer of $\varphi, \psi$. First observe that

$$(\varphi - \psi) \circ e = 0 \implies \varphi \circ e - \psi \circ e = 0 \implies \varphi \circ e = \psi \circ e$$

using bilinearity of $\circ$. Hence we see that $e$ equalizes $\varphi$ and $\psi$, although we now need to demonstrate its universal property.

Now suppose that there exists an object $K$ equipped with a morphism $\sigma : K \longrightarrow G$ such that $\psi \circ \sigma = \psi \circ \varphi$.

$$\begin{tikzcd}
\text{Ker}(\varphi - \psi) \arrow[r, "e"] & G \arrow[r, shift left, "\varphi"] \arrow[r, shift right, "0"'] & H \\
K \arrow[ur, "\sigma"'] & &
\end{tikzcd}$$

However, note that

$$\varphi \circ \sigma = \psi \circ \sigma \implies (\varphi - \psi) \circ \sigma = 0.$$

Since $e : \text{Ker}(\varphi) \longrightarrow G$ is kernel, we note that its universal property implies that because $(\varphi - \psi) \circ \sigma = 0$ that there must exists a unique morphism $u : K \longrightarrow \text{Ker}(\varphi)$ such that $e \circ u = \sigma$. Thus we have shown the diagram below

$$\begin{tikzcd}
\text{Ker}(\varphi - \psi) \arrow[r, "e"] & G \arrow[r, shift left, "\psi"] \arrow[r, shift right, "\varphi"'] & H \\
K \arrow[u, dashed, "u"] \arrow[ur, "\sigma"'] & &
\end{tikzcd}$$

must hold so that $(\text{Ker}(\varphi), e : \text{Ker}(\varphi) \longrightarrow G)$, is actually an equalizer! ∎

Note that we've once more utilized the bilinearity of $\circ$ to construct the above proof, which again reminds us that the assumptions we've made so far are necessary and useful. The above proof now motivates the following definition.

**Definition 8.3.5.** Let $\mathcal{C}$ be an additive category. Then we say $\mathcal{C}$ is **preabelian** if it has kernels and cokernels; or, equivalently, if it has all equalizers and coequalizers.

What we have on our hands now is a very nice category where (1) finite biproducts exist and (2) all equalizers and coequalizers exist. If we recall from our experience with limits, this automatically grants us the following proposition.

**Proposition 8.3.6.** A preabelian category has all finite limits and finite colimits.

*Proof.* If a category has finite products and equalizers, it has finite limits. If it has finite coproducts and coequalizers, it has finite colimits. This is Theorem 5.3.1. ■

The fact that there exist finite limits and colimits is extremely convenient in preabelian categories.

**Proposition 8.3.7.** Let $\mathcal{C}$ be a preabelian category. Let $J$ be a connected category and suppose $F : J \longrightarrow \mathcal{C}$ is a functor. Then

$$\operatorname{Lim} F \cong \operatorname{Colim} F.$$

*Proof.* Recall the limit satisfies universal property

$$\begin{tikzcd}
\Delta(\operatorname{Lim} F) \arrow[r, "u"] & F \\
\Delta(C) \arrow[u, dashed, "\Delta(h)"] \arrow[ur, "f"'] &
\end{tikzcd}
\quad \Longrightarrow \quad
\begin{tikzcd}
\operatorname{Lim} F \arrow[r, "u^i"] & F(i) \\
C \arrow[u, dashed, "h"] \arrow[ur, "f^i"'] &
\end{tikzcd}$$

for every object $C$ equipped with a family of morphisms $f^i : C \longrightarrow F(i)$. Construct the family of morphisms

$$f_i^j = \begin{cases} \emptyset_i^j : F(i) \longrightarrow F(j) & \text{if } i \neq j \\ 1_{F(i)} & \text{if } i = j \end{cases}$$

where $\emptyset_i^j : F(i) \longrightarrow F(j)$ is the unique zero morphism from $F(i)$ to $F(j)$. Then by the universal property of the limit, for each $i \in J$, there exists a unique morphism $h_i : F(i) \longrightarrow \operatorname{Lim} F$ such that the diagram below commutes.

$$\begin{tikzcd}
\operatorname{Lim} F \arrow[r, "u^j"] & F(j) \\
F(i) \arrow[u, dashed, "h_i"] \arrow[ur, "f_i^j"'] &
\end{tikzcd}$$

That is, we have $u^j \circ h_i = f_i^j$. We now argue that we have a colimit on our hands. Specifically, suppose $D$ is an object of $\mathcal{C}$ equipped with a family of morphisms $g_j : F(j) \longrightarrow D$. Then observe that we can supply a morphism

$$\sum_{k \in J} g_k u^k : \operatorname{Lim} F \longrightarrow D$$

where the addition operation is from the group structure of $\operatorname{Hom}(\operatorname{Lim} F, D)$, such that the diagram below commutes.

$$\begin{tikzcd}
\operatorname{Lim} F \arrow[d, dashed, "\sum_{k \in J} g_k u^k"'] & F(j) \arrow[l, "h_j"'] \arrow[dl, "g_j"] \\
D &
\end{tikzcd}$$

This diagram commutes since

$$\left(\sum_{k \in J} g_k u^k\right) \circ h_j = \sum_{k \in J} g_k (u^k \circ h_j) = g_j (u^j \circ h_j) = g_j$$

where we utilized the bilinearity of the composition operator. Thus we see that $\operatorname{Lim} F$ is behaving just like a colimit!

The only thing we must verify at this point is that this morphism is unique. Towards that goal, suppose that $\ell : \operatorname{Lim} F \longrightarrow D$ is another morphism such that $\ell \circ h_j = g_j$. Recall that $u^i \circ h_i = 1_{F(i)}$, so that $h_i$ is a monomorphism. Then observe that we can take the image of the map

$$h_i : F(i) \longrightarrow \operatorname{Lim} F$$

under the contravariant hom functor to get an epic group homomorphism

$$\operatorname{Hom}(\operatorname{Lim} F, D) \xrightarrow{\circ h_i} \operatorname{Hom}(F(i), D)$$

between abelian groups, as $\circ$ obeys bilinearity properties. By the first isomorphism theorem we then have that

$$\operatorname{Hom}(F(i), D) \cong \operatorname{Hom}(\operatorname{Lim} F, C) / \operatorname{Ker}(\circ h_i).$$

Now we want to show that this map is also injective, because then we could observe that since

$$\left(\ell - \sum_{k \in J} g_k \circ u^k\right) \circ h_i = 0$$

that

$$\ell - \sum_{k \in J} g_k \circ u^k = 0.$$

But it seems like we don't have enough to show that at the moment...

∎

# 8.4 Kernels and Cokernels

At this point we've discussed preadditive, additive, and preabelian categories, where preabelian categories are just additive categories with the additional hypothesis that kernels and cokernels exist. This additional hypothesis is extremely useful, so we will demonstrate what this implies for us.

Let $\mathcal{C}$ be a preabelian category. Consider an arbitrary morphism $f : A \longrightarrow B$. One way to think about kernels and Cokernels is that they give rise to objects in the comma categories $(\mathcal{C} \downarrow A)$ and $(B \downarrow \mathcal{C})$.

$$\begin{tikzcd} \operatorname{Ker}(f) \arrow[r, "e"] & A \arrow[r, shift left, "f"] \arrow[r, shift right, "0"'] & B \arrow[r, "c"] & \operatorname{Coker}(f) \end{tikzcd}$$

Now in the comma category $(\mathcal{C} \downarrow A)$, a morphism between two objects $(C, f : C \longrightarrow A)$ and $(D, g : D \longrightarrow A)$ is a morphism $h : D \longrightarrow C$ in $\mathcal{C}$ such that $f = g \circ h$. Similarly, a morphism in the comma category $(A \downarrow \mathcal{C})$ between two objects $(P, m : A \longrightarrow P)$ and $(Q, n : A \longrightarrow Q)$ is a morphism $k : P \longrightarrow Q$ such that $n = h \circ m$. These relations give rise to the bow-tie diagram:

$$\begin{tikzcd} C \arrow[dr, "f"] & & P \arrow[dd, "k"] \\ & A \arrow[ur, "m"] \arrow[dr, "n"'] & \\ D \arrow[uu, "h"] \arrow[ur, "g"'] & & Q \end{tikzcd}$$

With that said, we can actually turn these categories into partial orders. In $(\mathcal{C} \downarrow A)$, we say $g \leq f$ if there exists an $h$ such that $f \circ h = g$, and in $(A \downarrow \mathcal{C})$, we say $m \leq n$ if there exists a $k$ such that $n = k \circ m$.

It turns out that this perspective is actually quite useful.

**Proposition 8.4.1.** Let $\mathcal{C}$ be a category with a zero object, equalizers and coequalizers. Then for each object $A$ of $\mathcal{C}$, we have the functors

$$\operatorname{Ker} : (A \downarrow \mathcal{C}) \longrightarrow (\mathcal{C} \downarrow A)$$
$$\operatorname{Coker} : (\mathcal{C} \downarrow A) \longrightarrow (A \downarrow \mathcal{C}).$$

that assign kernels and cokernels. Moreover, these functors establish a antitone Galois correspondence; hence we have that

$$\operatorname{Ker}(\operatorname{Coker}(\operatorname{Ker}(f))) = \operatorname{Ker}(f) \quad \operatorname{Coker}(\operatorname{Ker}(\operatorname{Coker}(f))) = \operatorname{Coker}(f).$$

Therefore, any $\varphi$ is a kernel if and only if $\varphi = \operatorname{Ker}(\operatorname{Coker}(\varphi))$, while any $\psi$ is a cokernels if and only if $\psi = \operatorname{Coker}(\psi(\psi))$.

*Proof.* We demonstrate functoriality. First we want our functor to act on objects as

$$(C, f : A \longrightarrow C) \mapsto (\mathrm{Ker}(f), e_1 : \mathrm{Ker}(f) \longrightarrow A).$$

Now we explain how the functor works on morphisms. Suppose we have two objects of our comma category $(C, f : A \longrightarrow C)$ and $(D, g : A \longrightarrow D)$, and that $h : D \longrightarrow C$ is a morphism in $(A \downarrow \mathcal{C})$ from $(D, g : A \longrightarrow D)$ to $(C, f : A \longrightarrow C)$. Then we have the diagram below.

$$\begin{tikzcd}
\mathrm{Ker}(f) \arrow[dr, "e_1"] & & C \\
& A \arrow[ur, "f"] \arrow[dr, "g"'] & \\
\mathrm{Ker}(g) \arrow[ur, "e_2"'] & & D \arrow[uu, "h"']
\end{tikzcd}$$

Now note that

$$f \circ e_2 = (h \circ g) \circ e_2 = h \circ (g \circ e_2) = 0.$$

Thus, by the universal property of $e_1 : \mathrm{Ker}(f) \longrightarrow A$, we know there exists a *unique* morphism $h' : \mathrm{Ker}(g) \longrightarrow \mathrm{Ker}(f)$ such that the diagram below commutes.

$$\begin{tikzcd}
\mathrm{Ker}(f) \arrow[dr, "e_1"] & & C \\
& A \arrow[ur, "f"] \arrow[dr, "g"'] & \\
\mathrm{Ker}(g) \arrow[uu, "h'"] \arrow[ur, "e_2"'] & & D \arrow[uu, "h"']
\end{tikzcd}$$

However, this is exactly what it means to have a morphism between the objects $(\mathrm{Ker}(g), e_2 : \mathrm{Ker}(g) \longrightarrow A)$ and $(\mathrm{Ker}(f), e_1 : \mathrm{Ker}(f) \longrightarrow A)$. Thus we see that our functor maps on morphisms in $(A \downarrow \mathcal{C})$ in a nice way:

$$h \mapsto h' : (\mathrm{Ker}(g), e_2 : \mathrm{Ker}(g) \longrightarrow A) \longrightarrow (\mathrm{Ker}(f), e_1 : \mathrm{Ker}(f) \longrightarrow A).$$

where $h'$ is the unique map obtained from $h$ as explained above. With the remaining properties easily verified, this defines a functor between the categories. In addition, we can dualize our work above to also get the functor $\mathrm{Coker} : (\mathcal{C} \downarrow A) \longrightarrow (A \downarrow \mathcal{C})$.

Now this creates a Galois correspondence by regarding the comma categories as partially ordered sets. Suppose that $g \le \mathrm{Ker}(f)$. That is, there exists a $h$ such that $\mathrm{Ker}(f) \circ h = g$. Then we can compare $\mathrm{Coker}(g)$ and $f$ by considering the diagram below.

$$\begin{tikzcd}
\mathrm{Ker}(f) \arrow[dr, "e"] & & \mathrm{Coker}(g) \\
& A \arrow[ur, "c"] \arrow[dr, "f"'] & \\
B \arrow[uu, "h"] \arrow[ur, "g"'] & & C
\end{tikzcd}$$

Now observe that

$$f \circ g = f \circ (e \circ h) = 0 \circ h = 0.$$

Therefore, by the universal property of the cokernel, we know there exists a unique morphism $h' : \text{Coker}(g) \longrightarrow f$ such that the diagram below commutes. This then implies that $f \leq \text{Coker}(g)$.

$$\begin{tikzcd}
\text{Ker}(f) \arrow[drr, "e"] & & & & \text{Coker}(g) \arrow[dd, dashed, "h'"] \\
& & A \arrow[urr, "c"] \arrow[drr, "f"'] & & \\
B \arrow[uu, "h"] \arrow[urr, "g"'] & & & & C
\end{tikzcd}$$

By a similar argument, we have that if $f \leq \text{Coker}(g)$, then $g \leq \text{Ker}(f)$. Hence we have that

$$g \leq \text{Ker}(f) \iff f \leq \text{Coker}(g)$$

so that, as preorder, the kernel and cokernels functors are adjoint pairs that form an antitone Galois correspondence. Moreover, this implies that for each $f : B \longrightarrow A$ and $g : A \longrightarrow C$,

$$f \leq \text{Coker}(\text{Ker}(f)) \qquad g \leq \text{Ker}(\text{Coker}(g)).$$

In particular, if $f$ is the cokernel of some morphism $\varphi$, and if $g$ is the kernel of some morphism $\psi$, then we have that

$$\text{Coker}(\varphi) \leq \text{Coker}(\text{Ker}(\text{Coker}(\varphi))) \quad \text{Ker}(\psi) \leq \text{Ker}(\text{Coker}(\text{Ker}(\psi))).$$

However, applying the order reversing functors Coker and Ker on the relations $\varphi \leq \text{Ker}(\text{Coker}(\varphi))$ and $\psi \leq \text{Coker}(\text{Ker}(\psi))$ yields

$$\text{Coker}(\text{Ker}(\text{Coker}(\varphi))) \leq \text{Coker}(\varphi) \quad \text{Ker}(\text{Coker}(\text{Ker}(\psi))) \leq \text{Ker}(\psi).$$

Hence we have that $\text{Coker}(\text{Ker}(\text{Coker}(\varphi))) \cong \text{Coker}(\varphi)$ and $\text{Ker}(\text{Coker}(\text{Ker}(\psi))) \cong \text{Ker}(\psi)$ as desired. ∎

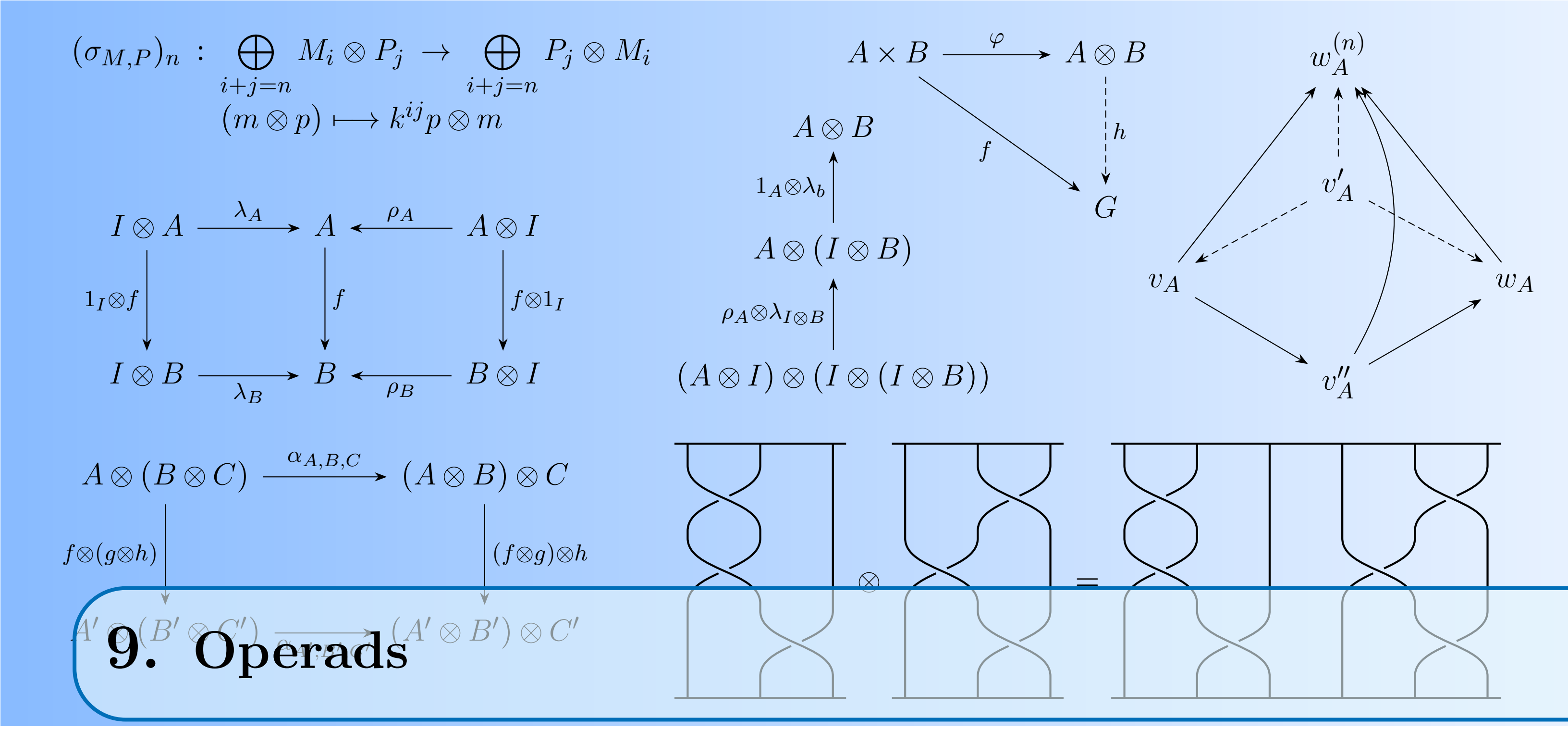
## 8.5 Abelian Categories

Let $\mathcal{C}$ be a preabelian category, and consider an arbitrary morphism $\varphi : A \longrightarrow B$. Then, since we are in an abelian category, we can calculate the kernel and cokernel of this morphism, which both have their familiar universal properties.

$$\begin{tikzcd}
\mathrm{Ker}(\varphi) \arrow[r, "e"] & A \arrow[r, shift left, "\varphi"] \arrow[r, shift right, "\emptyset"'] & B \arrow[r, "c"] \arrow[dr, "\psi"'] & \mathrm{Coker}(\varphi) \arrow[d, dashed, "k"] \\
C \arrow[u, dashed, "h"] \arrow[ur, "\varphi"'] & & & D
\end{tikzcd}$$

One thing we can do is examine both the kernel and the cokernel *of* these two morphisms. Specifically, we can calculate the kernel $\mathrm{Ker}(c)$ of $c$ and the cokernel $\mathrm{Coker}(e)$ of $e$. However, since we have a map $\varphi : A \longrightarrow B$ such that $c \circ \varphi = 0$, we see that there exists a unique map $u : A \longrightarrow \mathrm{Ker}(\mathrm{Coker}(f))$ such that $\varphi = e' \circ u$. Dually, since $\varphi \circ e = 0$, there exists a unique map $v : \mathrm{Coker}(\mathrm{Ker}(f)) \longrightarrow B$. such that $\varphi = v \circ c'$.

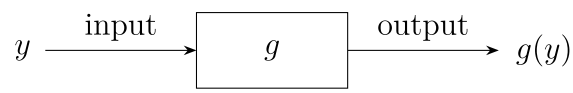
$$\begin{tikzcd}
& & \mathrm{Ker}(\mathrm{Coker}(f)) \arrow[dr, "e'"] & & \\
\mathrm{Ker}(f) \arrow[r, "e"] & A \arrow[ur, dashed, "u"] \arrow[rr, "\varphi"] \arrow[dr, "c'"'] & & B \arrow[r, "c"] \arrow[ddr, "\psi"'] & \mathrm{Coker}(f) \arrow[dd, dashed, "k"] \\
& & \mathrm{Coker}(\mathrm{Ker}(f)) \arrow[ur, dashed, "v"'] & & \\
C \arrow[uu, dashed, "h"] \arrow[uur, "\varphi"'] & & & & D
\end{tikzcd}$$



# 9. Operads

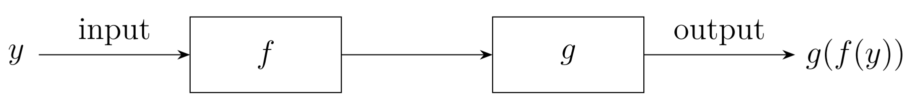
## 9.1 Operads on Sets

Let $Y, Z$ be sets. Consider a function $g : Y \longrightarrow Z$. The way we've been taught to think about this function is as a process where we're sending an element $y \mapsto g(y)$ in a well-defined manner.




*The typical picture one uses when describing a function.*

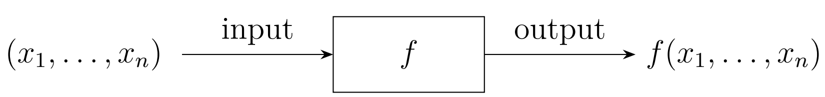
Furthermore, if we have another function $f : X \longrightarrow Y$, then we can set up a pipeline $x \mapsto f(x) \mapsto g(f(x))$. This then establishes an obvious function $g \circ f : X \longrightarrow Z$.




*The function $g \circ f$.*

But the way that we've thought about functions, and more generally morphisms, is actually over-simplistic. Here we will demonstrate that we can *generalize the concept of morphism composition.*

Denote $\text{End}_n(X)$ to be the set of all functions $f : X^n \longrightarrow X$. Then for such a function, if we stick with our simplistic concept of plugging things in, we imagine something like




However, a more natural way is to imagine that we're taking values $n$-many values $x_i \in X$ and plugging them into the function $f : X^n \longrightarrow X$. That is, we don't have to just think of *one* $g : Y \longrightarrow X^n$ to form a concept of composition. We can instead imagine that each of these $x_i$ values came from functions $g_1 : Y_1 \longrightarrow X$, $g_2 : Y_2 \longrightarrow X, \cdots, g_n : Y_n \longrightarrow X$.

$$\begin{tikzcd}
Y_1 \arrow[dr, "g_1"'] & Y_2 \arrow[d, "g_2"] & \cdots & Y_n \arrow[dll, "g_n"] \\
& X \times X \times \cdots \times X \arrow[d, "f"] & & \\
& X & &
\end{tikzcd}$$

This is in its own right a function; a function from $Y_1 \times Y_2 \times Y_n \longrightarrow X$. It's a generalization of function composition; when we only have one $g_1$ we just get back our original notion of function composition. We've been restricting ourselves this whole time. Now to make this even more interesting, suppose $Y_1 = X^{a_1}, Y_2 = X^{a_2}, \dots, Y_n = X^{a_n}$ where $a_1, a_2, \dots, a_n$ are positive integers. That is, suppose we have that $g_i \in \text{End}_{a_i}(X)$.

$$\begin{tikzcd}
X^{a_1} \arrow[dr, "g_1"'] & X^{a_2} \arrow[d, "g_2"] & \cdots & X^{a_n} \arrow[dll, "g_n"] \\
& X \times X \times \cdots \times X \arrow[d, "f"] & & \\
& X & &
\end{tikzcd}$$

The above composition can be expressed as $f(g_1, g_2, \dots, g_n)$ which we may denote as

$$f \circ_{a_1, a_2, \dots, a_n} (g_1, g_2, \dots, g_n) : X^{a_1} \times X^{a_2} \times \cdots \times X^{a_n} \longrightarrow X.$$

and note that we've construction a function in $\text{End}_{a_1+a_2+\cdots+a_n}(X)$ using one $f \in \text{End}_n(X)$ and $n$-many $g_i \in \text{End}_i(X)$. Then what we see is that our composition map is really a function that can be written formally as

$$\circ_{a_1, a_2, \dots, a_n} : X^n \times (X^{a_1} \times X^{a_2} \times \cdots \times X^{a_n}) \longrightarrow X^{a_1+a_2+\cdots+a_n}$$

Then we can make this even more interesting. Each $g_i : X^{a_i} \longrightarrow X$ is *just like* $f : X^n \longrightarrow X$. Hence we can repeat the same process on each $g_i$, and plug a family of functions $h_{i,j} : X^{k_{i,j}} \longrightarrow X$ where $j = 1, 2, \dots, a_i$.

$$\begin{tikzcd}
X^{k_{1,1}}, \; X^{k_{1,2}}, \dots, \; X^{k_{1,a_1}} \arrow[d, "h_{1,1}"', "h_{1,2}", "h_{1,a_1}"] & X^{k_{2,1}}, \; X^{k_{2,2}}, \dots, \; X^{k_{2,a_2}} \arrow[d, "h_{2,1}"', "h_{2,2}", "h_{2,a_2}"] & \cdots & X^{k_{n,1}}, \; X^{k_{n,2}}, \dots, \; X^{k_{1,a_n}} \arrow[d, "h_{n,1}"', "h_{n,2}", "h_{n,a_n}"] \\
X \times X \times \overbrace{\cdots}^{a_1\text{-times}} \times X \arrow[dr, "g_1"'] & X \times X \times \overbrace{\cdots}^{a_2\text{-times}} \times X \arrow[d, "g_2"] & \cdots & X \times X \times \overbrace{\cdots}^{a_n\text{-times}} \times X \arrow[dll, "g_n"] \\
& X \times X \times \cdots \times X \arrow[d, "f"] & & \\
& X & &
\end{tikzcd}$$

Now there are two ways to think about this function. There is

$$[f \circ_{a_1,a_2,\dots,a_n} (g_1, g_2, \dots, g_n)] \circ_{k_{1,1},\dots,k_{1,a_1},\dots,k_{n,a_1},\dots,k_{n,a_n}} (h_{1,1}, \dots, h_{n,a_n})$$

which first composes $f$ with the $g$-family, and then composes with the $h$-family, and then there is

$$f \circ_{(k_{1,1}+\cdots+k_{1,a_1}),\dots,(k_{n,1}+\cdots+k_{n,a_n})} \Big(g_1 \circ_{k_{1,1},\dots,k_{1,a_1}} (h_{1,1}, \dots, h_{1,a_1}), \dots, g_n \circ_{k_{n,1},\dots,k_{n,a_n}} (h_{n,1}, \dots, h_{n,a_n})\Big)$$

which first composes each $g$ with its respective $h$-family, and then composing the resulting structure with $f$. Since these are just functions, and individual composition is associative, the above two ways are the same. This construction which we have demonstrated is an example of an *operad*; specifically, a symmetric operad. The previous example can now be seen as motivation for the following two definitions (which will definitely need repeated read-overs).

**Definition 9.1.1.** A **nonsymmetric operad** $X$ in **Set** consists of a family of sets $\{X_n\}_{n=1}^{\infty}$, an identity element $I \in X_1$ (whose purpose will soon be elaborated), and a composition map

$$\begin{aligned} \circ_{n,a_1,a_2,\dots,a_n} : X_n \times (X_{a_1} \times X_{a_2} \times \cdots \times X_{a_n}) &\longrightarrow X_{a_1+a_2+\cdots+a_n} \\ (f, g_1, g_2, \dots, g_n) &\mapsto f \circ_{a_1,a_2,\dots,a_n} (g_1, g_2, \dots, g_n) \end{aligned}$$

which must exist for each $n = 1, 2, \dots,$ and any $a_1, a_2, \dots, a_n \in \mathbb{N}$, such that

**(NS-OP1: Associativity.)** Let $n \in \mathbb{N}$ and consider $f \in X_n$. Let $a_1, a_2 \dots, a_n \in \mathbb{N}$. Then

$$\begin{gathered} [f \circ_{a_1,a_2,\dots,a_n} (g_1, g_2, \dots, g_n)] \circ_{k_{1,1},\dots,k_{1,a_1},\dots,k_{n,a_1},\dots,k_{n,a_n}} (h_{1,1}, \dots, h_{n,a_n}) \\ = \\ f \circ_{(k_{1,1}+\cdots+k_{1,a_1}),\dots,(k_{n,1}+\cdots+k_{n,a_n})} \Big(g_1 \circ_{k_{1,1},\dots,k_{1,a_1}} (h_{1,1}, \dots, h_{1,a_1}), \dots, g_n \circ_{k_{n,1},\dots,k_{n,a_n}} (h_{n,1}, \dots, h_{n,a_n})\Big) \end{gathered}$$

**(NS-OP2): Identity.** For every $f \in X_n$ we have that

$$f \circ_{1,1,\dots,1} (I, I, \dots, I) = f = I \circ_n (f).$$

**Definition 9.1.2.** A **symmetric operad** is a nonsymmetric operad $X$ with a right group action $\cdot_n : X_n \times S_n \longrightarrow X_n$ by the symmetric group $S_n$ for each $n = 1, 2, \dots,$ subject to the following axioms.

**(S-OP1: Equivariance 1)** Let $f \in X_n$ and pick $g_1 \in X_{a_1}, \dots, g_n \in X_{a_n}$ for some $a_1, a_2, \dots, a_n \in \mathbb{N}$. Then for a $\tau \in S_n$, we must have

$$(f \cdot \tau) \circ_{a_1,\dots,a_n} (g_1, \dots, g_n) = (f \circ_{a_{\tau^{-1}(1)},\dots,a_{\tau^{-1}(n)}} (g_{\tau^{-1}(1)}, \dots, g_{\tau^{-1}(n)})) \cdot \tau'$$

where $\tau' \in S_{a_1+\cdots+a_n}$. Here, $\tau'$ is a *block permutation* that swaps the $i$-th block with the $\tau(i)$-th block. That is, if $\tau \in S^n$ as a permutation acts as

$$(1, 2, \dots, n) \mapsto (\tau(1), \tau(2), \dots, \tau(n))$$

then $\tau' \in S_{a_1+a_2+\cdots+a_n}$ acts as

$$(\overbrace{1,2,\dots,a_1}^{\text{1st block}},\dots,\overbrace{a_1+\cdots+a_i+1,\dots a_1+\cdots+a_{i+1}}^{i\text{-th block}},\dots\overbrace{a_1+\cdots+a_{n-1}+1,\dots,a_1+\cdots+a_n}^{n\text{-th block}})$$

$$\mapsto$$

$$(\dots,\overbrace{1,2,\dots,a_1}^{\tau(1)\text{-th block}},\dots,\overbrace{a_1+\cdots+a_i+1,\dots,a_1+\cdots+a_{i+1}}^{\tau(i)\text{-th block}},\dots,\overbrace{a_1+\cdots+a_{n-1}+1,\dots,a_1+\cdots+a_n}^{\tau(n)\text{-th block}},\dots$$

**(S-OP2: Equivariance 2)** Let $f, g_i$ is as above, and choose $\sigma_1 \in S_1, \dots, \sigma_n \in S_n$. Then we have that

$$f \circ_{a_1,\dots,a_n} (g_1 \cdot \sigma_1, \dots, g_n \cdot \sigma_n) = (f \circ_{a_1,\dots,a_n} (g_1, \dots, g_n)) \cdot (\sigma_1, \dots, \sigma_n)$$

where $(\sigma_1, \sigma_2, \dots, \sigma_n) \in S_{a_1+a_2+\cdots+a_n}$ is the permutation described as below.

$$(\overbrace{1,2,\dots,a_1}^{\text{1st block}},\dots,\overbrace{a_1+\cdots+a_{n-1}+1,,\dots,a_1+\cdots+a_{n-1}a_n}^{n\text{-th block}})$$

$$\mapsto$$

$$(\underbrace{\sigma_1(1),\sigma_1(2),\dots,\sigma_1(a_1)}_{\text{1st block}},\dots,\underbrace{a_1+\cdots+a_{n-1}+\sigma_n(1),\dots,a_1+\cdots+a_{n-1}+\sigma_n(a_n)}_{n\text{-th block}})$$

---

**Example 9.1.3.** We can continue with our previous construction concerning the family of sets

$$\text{End}_n(X) = \{f : X^n \longrightarrow X \mid f \in \textbf{Set}\}$$

to demonstrate that it forms a symmetric operad. As we already established associativity **NS-OP1**, we need to verify the identity axiom **NS-OP2**. Such an identity element can be chosen if we select $I = 1_X : X \longrightarrow X$. On one hand we have for any $f \in X^n$ that

$$f \circ_{1,1,\dots,1} (I, I, \dots, I) = f(1_x, 1_x, \dots, 1_x) = f$$

while on the other we have that $I \circ_n f = 1_X \circ f = f$. Next, define a group action of $S_n$ on $\text{End}_n(X)$ as

$$(f \cdot \sigma)(x_1, x_2, \dots, x_n) = f(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}).$$

We now verify **S-OP1** with this group action. Let $f \in \text{End}_n(X)$ and $g_i \in \text{End}_i(X)$ for $i = 1, 2, \dots, n$. For a given $\tau \in S_n$, consider the points $(x_1, \dots, a_1) \in X^{a_1}, \dots, (x_1, \dots, a_n) \in X^{a_n}$. Observe that $(f \cdot \tau) \circ_{a_1,\dots,a_n} (g_1, \dots, g_n)$ first plugs in the each $(x_{a_i-1}, \dots, x_{a_i})$ into $g_i$, which is then plugged into $f$. However, the action of $\tau$ swaps these resulting coordinates. Thus we get that

$$(f \cdot \tau) \circ_{a_1,\dots,a_n} (g_1, \dots, g_n)(x_1, \dots, a_1, \dots, x_{a_{n-1}+1}, \dots, x_{a_n}) = (\dots, \overbrace{g_i(x_{a_{i-1}+1}, \dots, x_{a_i})}^{\tau(i)-\text{th entry}}, \dots)$$

How do we write this more formally? Well, to answer that, we need to know the answer to the following question: which $g_i(x_{a_{i-1}+1}, \dots, x_{a_i})$ maps to, say, the 1st coordinate? This is equivalently to asking: what is $\tau^{-1}(1)$? Hence we see that

$$
\begin{aligned}
&(f \cdot \tau) \circ_{a_1,\dots,a_n} (g_1, \dots, g_n)(x_1, \dots, a_1, \dots, x_{a_{n-1}+1}, \dots, x_{a_n})\\
&= f(g_{\tau^{-1}(1)}(x_{a_{\tau^{-1}(1)-1}+1}, \dots, x_{a_{\tau^{-1}(1)}}), \dots, g_{\tau^{-1}(n)}(x_{a_{\tau^{-1}(n)-1}+1}, \dots, x_{a_{\tau^{-1}(n)}}))\\
&= f \circ_{\tau^{-1}(1),\tau^{-1}(2),\dots,\tau^{-1}(n)} (g_{\tau^{-1}(1)}, g_{\tau^{-1}(2)}, \dots, g_{\tau^{-1}(n)})(x_{a_{\tau^{-1}(1)-1}+1}, \dots, x_{a_{\tau^{-1}(1)}}, \dots, x_{a_{\tau^{-1}(n)-1}+1}, \dots, x_{a_{\tau^{-1}(n)}})\\
&= \Big(f \circ_{\tau^{-1}(1),\tau^{-1}(2),\dots,\tau^{-1}(n)} (g_{\tau^{-1}(1)}, g_{\tau^{-1}(2)}, \dots, g_{\tau^{-1}(n)}) \cdot \tau'\Big)(x_1, \dots, x_{a_1}, \dots, x_{a_{n-1}+1}, \dots, x_{a_n}).
\end{aligned}
$$

where $\tau'$ is the block permutation described in the definition. Thus we see that

$$
(f \cdot \tau) \circ_{a_1,\dots,a_n} (g_1, \dots, g_n) = f \circ_{\tau^{-1}(1),\tau^{-1}(2),\dots,\tau^{-1}(n)} (g_{\tau^{-1}(1)}, g_{\tau^{-1}(2)}, \dots, g_{\tau^{-1}(n)}) \cdot \tau'
$$

as desired. Thus we have **S-OP1**. Finally, we show **S-OP2**, which is a bit easier to demonstrate. As before, let $f, a_i$ and $g_i$ be as described before. Let $\sigma_1 \in S_1, \dots, \sigma_n \in S_n$. Then

$$
\begin{aligned}
&f \circ_{a_1,\dots,a_n} (g_1 \cdot \sigma_1, \dots, g_n \cdot \sigma_n)(x_1, \dots, x_{a_1}, \dots, x_{a_{n-1}+1}, \dots, x_{a_n})\\
&= f\Big((g_1 \cdot \sigma_1)(x_1, \dots, x_{a_1}), \dots, (g_n \cdot \sigma_n)(x_{a_{n-1}+1}, \dots, x_{a_n})\Big)\\
&= f\Big(g_1(x_{\sigma_1(1)}, \dots, x_{\sigma_1(a_1)}), \dots, g_n(x_{\sigma_n(1)}, \dots, x_{\sigma_n(a_n)})\Big)\\
&= \Big(f \circ_{a_1,\dots,a_n} (g_1, \dots, g_n)\Big)(x_{\sigma_1(1)}, \dots, x_{\sigma_1(a_1)}, \dots, x_{\sigma_n(1)}, \dots, x_{\sigma_n(a_n)})\\
&= \Big(f \circ_{a_1,\dots,a_n} (g_1, \dots, g_n)\Big) \cdot (\sigma_1, \dots, \sigma_n)(x_1, \dots, x_{a_1}, \dots, x_{a_{n-1}+1}, \dots, x_{a_n})
\end{aligned}
$$

Thus we see that

$$
f \circ_{a_1,\dots,a_n} (g_1 \cdot \sigma_1, \dots, g_n \cdot \sigma_n) = (f \circ_{a_1,\dots,a_n} (g_1, \dots, g_n)) \cdot (\sigma_1, \dots, \sigma_n)
$$

so that **S-OP2** is satisfied. All together, we have that for any set $X$, the family of sets $\text{End}_n(X)$ forms a symmetric operad.

---

---

**Example 9.1.4.** Consider the family of sets $\text{Assoc}_n = S_n$ where each level is the $n$-th symmetric group. Suppose that $\tau \in S_n$ and that $\sigma_1 \in S_{a_1}, \sigma_2 \in S_{a_2}, \dots, \sigma_n \in S_{a_n}$ for $a_1, a_2, \dots, a_n \in \mathbb{N}$. Then we define

$$
\tau \circ_{a_1,\dots,a_n} (\sigma_1, \sigma_2, \dots, \sigma_n) \in S_{a_1+a_2+\cdots+a_n}
$$

as a permutation of $a_1+a_2+\cdots+a_n$ letters. Before we describe the permutation, we'll introduce some notation. Consider the (ordered) tuple of the first $a_1 + \cdots + a_n$ integers.

$$(1, 2, \dots, a_1, a_1+1, a_1+2, \dots, a_1+a_2, \dots (a_1+\cdots+a_{n-1})+1, \dots, (a_1+\cdots+a_{n-1})+a_n)$$

We can more compactly denote this tuple as

$$(1, 2, \dots, a_1, 1', 2', \dots, a_2', \dots, 1', 2', \dots, a_n')$$

where from either context or coloring it will be clear what each $1', 2', \dots$ indicates. For example, above we'll have that $1' = a_1 + 1$ and $2' = a_1 + 2$ wheres $1' = (a_1 + \cdots + a_{n-1}) + 1$ and $2' = (a_1 + \cdots + a_{n-1}) + 2$. With that said, we can define $\tau \circ_{a_1, \dots, a_n} (\sigma_1, \sigma_2, \dots, \sigma_n) \in S_{a_1 + a_2 + \cdots + a_n}$ by its action on such a tuple, pictured below.

$$(1, 2, \dots, a_1, 1', 2', \dots, a_2', \dots, 1', 2', \dots, a_n')$$

$$\downarrow \sigma_1 \qquad \downarrow \sigma_2 \qquad \cdots \qquad \downarrow \sigma_n$$

$$(\underbrace{\sigma_1(1), \sigma_1(2)), \dots \sigma_1(a_1)}_{\text{1st block}}, \underbrace{\sigma_1'(1), \sigma_1'(2)), \dots \sigma_1'(a_1)}_{\text{2nd block}}, \dots, \underbrace{\sigma_n'(1), \sigma_n'(2) \dots, \sigma_n'(a_n)}_{a_n\text{-th block}})$$

$$\downarrow \tau$$

$$(\dots, \underbrace{\sigma_1(1), \sigma_1(2)), \dots \sigma_1(a_1)}_{\tau(1)\text{-th block}}, \dots, \underbrace{\sigma_1'(1), \sigma_1'(2)), \dots \sigma_1'(a_1)}_{\tau(2)\text{-th block}}, \dots, \underbrace{\sigma_n'(1), \sigma_n'(2), \dots, \sigma_n'(a_n)}_{\tau(a_n)\text{-th block}}, \dots, )$$

which can be rewritten more formally as

$$(\overbrace{\sigma'_{\tau^{-1}(1)}(1), \sigma'_{\tau^{-1}(1)}(2), \dots, \sigma'_{\tau^{-1}(1)}(a_{\tau^{-1}(1)})}^{\text{1st block}}, \dots, \overbrace{\sigma'_{\tau^{-1}(n)}(1), \sigma'_{\tau^{-1}n}(2), \dots, \sigma'_{\tau^{-1}(n)}(a_{\tau^{-1}(n)})}^{n\text{-th block}}) \in S_{a_1 + \cdots + a_n}.$$

Now for each $\sigma_i \in S_{a_i}$, let $\rho_{i,j} \in S_{k_{i,j}}$ for $j = 1, 2, \dots, a_i$ and for $k_{i,j} \in \mathbb{N}$. For notational convenience, denote $K = k_{1,1} + \cdots + k_{1,a_1} + \cdots + k_{n,1} + \cdots + k_{n,a_n}$. By our above definition, we can construct a permutation in $S_K$ by composing $\tau$ with the $\sigma$-family and with the $\rho$-family. There are two possible ways to construct such a permutation (and we'll show that they are equivalent, therefore satisfying **NS-OP1**). But before we do that we must consider the first $K$ integers. This will be a *huge* tuple; in full notation this is

$$\Big(\overbrace{1, 2, \dots, k_{1,1}}^{\text{1st block}}, \overbrace{k_{1,1}+1, k_{1,1}+2, \dots, k_{1,1}+k_{1,2}}^{\text{2nd block}}, \dots$$

$$\dots \overbrace{(k_{1,1}+k_{1,2}+\cdots+k_{1,a_1-1})+1, (k_{1,1}+k_{1,2}+\cdots+k_{1,a_1-1})+2, \dots, (k_{1,1}+k_{1,2}+\cdots+k_{1,a_1-1})+k_{1,a_1}}^{a_1\text{-th block}} \dots$$

$$\dots \overbrace{\sum_{i=1}^{n-1}\sum_{j=1}^{a_i} k_{i,j}+1 \sum_{i}^{n-1}\sum_{j=1}^{a_i} k_{i,j}+2, \dots, \sum_{i}^{n-1}\sum_{j=1}^{a_i} k_{i,j}+k_{n,1}}^{(a_1+a_2+\cdots+a_{n-1}+1)\text{-th block}}, \dots$$

$$\dots, \overbrace{\sum_{i}^{n-1}\sum_{j=1}^{a_i} k_{i,j}+(k_{n,1}+\cdots+k_{n,(a_n-1)})+1, \sum_{i}^{n-1}\sum_{j=1}^{a_i} k_{i,j}+(k_{n,1}+\cdots+k_{n,(a_n-1)})+2, \dots, \sum_{i}^{n}\sum_{j=1}^{a_i} k_{i,j}}^{(a_1+a_2+\cdots+a_{n-1}+a_n)\text{-th block}}\Big)$$

Using our previous notation we can rewrite this as

$$(\overbrace{1, 2, \dots, k_{1,1}}^{\text{1st block}}, \overbrace{1', 2', \dots, k_{1,2}}^{\text{2nd block}}, \dots, \overbrace{1', 2', \dots, k_{1,a_1}}^{a_1\text{-th block}}, \dots, \overbrace{1', 2', \dots, k_{n,1}}^{(a_1+\cdots+a_{n-1}+1)\text{-th block}}, \dots, \overbrace{1', 2', \dots, k_{n,a_n}}^{(a_1+\cdots+a_n)\text{-th block}})$$

where again, for example, $1' = k_{1,1}+1$ whereas $1' = \sum_{i}^{n-1}\sum_{j=1}^{a_i} k_{i,j}+(k_{n,1}+\cdots+k_{n,(a_n-1)})+1$.

Now we will first want to calculate

$$(\tau \circ_{a_1,\dots,a_n} (\sigma_1, \sigma_2, \dots, \sigma_n)) \circ_{k_{1,1},\dots,k_{1,a_1},\dots,k_{n,1},\dots,k_{n,a_n}} \circ(\rho_{1,1}, \dots, \rho_{n,a_n}).$$

The first step to computing this is to note that each $\rho_{i,j}$ permutes the numbers *within its block*.

$$\begin{array}{c}
(\overbrace{1, 2, \dots, k_{1,1}}^{\text{1st block}}, \overbrace{1', 2', \dots, k_{1,2}}^{\text{2nd block}}, \dots, \overbrace{1', 2', \dots, k_{i,j}}^{(a_1+\cdots+a_{i-1}+j)\text{-th block}}, \dots, \overbrace{1', 2', \dots, k_{n,a_n}}^{(a_1+\cdots+a_n)\text{-th block}}) \\
\downarrow \rho_{1,1} \qquad \cdots \qquad \downarrow \rho_{i,j} \qquad \cdots \qquad \downarrow \rho_{n,a_n} \\
(\underbrace{\rho_{1,1}(1), \rho_{1,1}(2), \dots, \rho_{1,1}(k_{1,1})}_{\text{1st block}}, \dots, \underbrace{\rho'_{i,j}(1)\rho'_{i,j}(2), \dots, \rho'_{i,j}(k_{i,j})}_{(a_1+\cdots+a_{i-1}+j)\text{-th block}}, \dots, \underbrace{\rho'_{n,a_n}(1), \rho'_{n,a_n}(2), \dots, \rho'_{n,a_n}(k_n, a_n)}_{(a_1+\cdots+a_n)\text{-th block}}))
\end{array}$$

Now that we've applied the $\rho$ permutations, we must apply the permutation $\tau \circ_{a_1,\dots,a_n} (\sigma_1, \sigma_2, \dots, \sigma_n)$ in $S_{a_1+\cdots+a_n}$. This will instead be a block permutation. Hopefully it is now clear why we were paying so much attention and to and keeping track of the blocks; we knew ahead of time that we were going to permute our $a_1 + \cdots + a_n$ blocks by using our $S_{a_1+\cdot+a_n}$ permutation $\tau \circ_{a_1,\dots,a_n} (\sigma_1, \sigma_2, \dots, \sigma_n)$ in $S_{a_1+\cdots+a_n}$.

Recall that for $\rho_{i,j}$, $i$ ranges from 1 to $n$ while $j$ ranges from 1 to $a_i$. Hence if we permute a block, we can represent it as follows.

$$(\underbrace{\rho_{1,1}(1),\rho_{1,1}(2),\dots,\rho_{1,1}(k_{1,1})}_{\text{1st block}},\dots,\underbrace{\rho'_{i,j}(1)\rho'_{i,j}(2),\dots,\rho'_{i,j}(k_{i,j})}_{(a_1+\cdots+a_{i-1}+j)\text{-th block}},\dots,\underbrace{\rho'_{n,a_n}(1),\rho'_{n,a_n}(2),\dots,\rho'_{n,a_n}(k_{n,a_n})}_{(a_1+\cdots+a_n)\text{-th block}}))$$

$$(\dots,\underbrace{\rho_{1,1}(1),\rho_{1,1}(2),\dots,\rho_{1,1}(k_{1,1})}_{\tau\circ_{a_1,\dots,a_n}(\sigma_1,\sigma_2,\dots,\sigma_n)(1)\text{th block}},\dots,\underbrace{\rho'_{n,a_n}(1),\rho'_{n,a_n}(2),\dots,\rho'_{n,a_n}(k_n,a_n)}_{\tau\circ_{a_1,\dots,a_n}(\sigma_1,\sigma_2,\dots,\sigma_n)(a_1+\cdots+a_n)\text{-th block}},\dots,)$$

which can be written more formally (that is, more horribly) as

$$(\dots,\underbrace{\rho_{\tau^{-1}(i),\sigma^{-1}_{\tau^{-1}(i)}(j)}(1),\rho_{\tau^{-1}(i),\sigma^{-1}_{\tau^{-1}(i)}(j)}(2),\dots,\rho_{\tau^{-1}(i),\sigma^{-1}_{\tau^{-1}(i)}(j)}(k_{\tau^{-1}(i),\sigma^{-1}_{\tau^{-1}(i)}(j)})}_{(a_1+\cdots+a_{i-1}+j)\text{-th block}},\dots)$$

At this point we'll want to see that this is the same as

$$\tau\circ_{(k_{1,1}+\cdots+k_{1,a_1}),\dots,(h_{n,1}+\cdots+k_{n,a_n})}(\sigma_1\circ_{k_{1,1},\dots,k_{1,a_1}}(\rho_{1,1},\dots,\rho_{1,a_1}),\dots,\sigma_n\circ_{k_{n,1},\dots,k_{n,a_n}}(\rho_{n,1},\dots,\rho_{n,a_n}))$$

To do this we need to think about each $\sigma_i\circ_{k_{i,1},\dots,k_{i,a_i}}(\rho_{i,1},\dots,\rho_{i,a_i})$ which isn't too bad. Each is a permutation in $S_{k_{i,1}+\cdots+k_{i,a_i}}$, and hence a permutation of the (ordered) tuple below.

$$(1,2,\dots,k_{i,1},k_{i,1}+1,k_{i,1}+2,\dots,k_{i,1}+k_{i,2},\dots,(k_{i,1}+k_{i,2}+\cdots+k_{i,a_{i-1}})+1,\dots,(k_{i,1}+k_{i,2}+\cdots)+k_{i,a_i})$$

which we again abbreviate as

$$(1,2,\dots,k_{i,1},1',2',k_{i,2},\dots,1',2',\dots,k_{i,a_i}).$$

With those notation above each permutation acts as

$$(\overbrace{1,2,\dots,k_{i,1}}^{\text{1st block}},\overbrace{1',2',k_{i,2}}^{\text{2nd block}},\dots,\overbrace{1',2',\dots,k_{i,a_i}}^{a_i\text{-th block}})$$

$$(\dots,\overbrace{\rho_{i,1}(1),\rho_{i,1}(2),\dots,\rho_{i,1}(k_{i,1})}^{\sigma_i(1)\text{-th block}},\dots,\overbrace{\rho'_{i,a_i}(1),\rho'_{i,a_i}(2),\dots,\rho'_{i,a_i}(k_{i,a_i})}^{\sigma_i(a_i)\text{-th block}},\dots)$$

which can be more formally understood as the tuple

$$(\overbrace{\rho'_{i,\sigma_i^{-1}(1)}(1),\rho'_{i,\sigma_i^{-1}(1)}(2),\dots,\rho'_{i,\sigma_i^{-1}(1)}(k_{i,\sigma_1^{-1}(1)})}^{\text{1st tuple}},\dots,\overbrace{\rho'_{i,\sigma_i^{-1}(a_i)}(1),\rho'_{i,\sigma_i^{-1}(a_i)}(2),\dots,\rho'_{i,\sigma_i^{-1}(a_i)}(k_{i,\sigma_i^{-1}(a_i)})}^{a_i\text{-th tuple}}) \tag{9.1}$$

Now that we understand what each $\sigma_i\circ_{k_{i,1},\dots,k_{i,a_i}}(\rho_{i,1},\dots,\rho_{i,a_i})$ does for $i=1,2,\dots,n$, and because we know that $\tau\in S_n$, this means we can compose $\tau$ with this family of $n$-permutations, which will give rise to a $S_{k_{1,1}+\cdots+k_{1,a_1}+\cdots+k_{n,1}+\cdots+k_{n,a_n}}$ permutation. To calculate this we just now directly apply their composition. This will act on the $k_{1,1}+\cdots k_{1,a_1}+\cdots+k_{n,1}+\cdots+k_{n,a_n}$ tuple

$$\underbrace{(\overbrace{1,2,\dots,k_{1,1}}^{\text{1st block}},\overbrace{1',2',k_{1,2}}^{\text{2nd block}},\dots,\overbrace{1',2',\dots,k_{1,a_1}}^{a_1\text{-th block}}}_{\text{1st } n\text{-block}},\dots,\underbrace{\overbrace{1,2,\dots,k_{n,1}}^{(a_1+\cdots+a_{n-1}+1)\text{-th block}},\underbrace{1',2',\dots,k_{n,2}}_{(a_1+\cdots+a_{n-1}+2)\text{-th block}},\dots,\overbrace{1',2',\dots,k_{i,a_i}}^{(a_1+\cdots+a_{n-1}+a_n)\text{-th block}})}_{n\text{-th } n\text{-block}}$$

by rearranging the tuple as below

$$(\dots,\underbrace{\dots,\overbrace{\rho_{1,1}(1),\rho_{1,1}(2),\dots,\rho_{1,1}(k_{1,1})}^{\sigma_1(1)\text{-block}},\dots,\overbrace{\rho'_{1,a_1}(1),\rho'_{1,a_1}(2),\dots,\rho'_{1,a_1}(k_{1,a_1})}^{\sigma_1(a_1)\text{-block}},\dots}_{\tau(1)\text{-th } n\text{-block}}$$

$$\dots,\underbrace{\dots,\overbrace{\rho_{n,1}(1),\rho_{n,1}(2),\dots,\rho_{n,1}(k_{n,1})}^{\sigma_n(1)\text{-block}},\dots,\overbrace{\rho_{n,a_n}(1),\rho_{n,a_n}(2),\dots,\rho_{n,a_n}(k_{n,a_n})}^{\sigma_n(a_n)\text{-block}},\dots}_{\tau(n)\text{-th } n\text{-block}})$$

and using (9.1) we know that this becomes

$$(\dots,\underbrace{\overbrace{\rho_{1,\sigma_1^{-1}(1)}(1),\rho_{1,\sigma_1^{-1}(1)}(2),\dots,\rho_{1,\sigma_1^{-1}(1)}(k_{1,\sigma_1^{-1}(1)})}^{(a_1+\cdots+a_{\tau(1)-1}+1)\text{-th tuple}},\dots,\overbrace{\rho_{1,\sigma_1^{-1}(a_1)}(1),\rho_{1,\sigma_1^{-1}(a_1)}(2),\dots,\rho_{1,\sigma_1^{-1}(a_1)}(k_{1,\sigma_1^{-1}(a_1)})}^{(a_1+\cdots+a_{\tau(1)-1}+a_{\tau(1)})\text{-th tuple}}}_{\tau(1)\text{-th } n\text{-block}}$$

$$\dots,\underbrace{\overbrace{\rho_{n,\sigma_n^{-1}(1)}(1),\rho_{n,\sigma_n^{-1}(1)}(2),\dots,\rho_{n,\sigma_n^{-1}(1)}(k_{n,\sigma_n^{-1}(1)})}^{(a_1+\cdots+a_{\tau(1)-1}+1)\text{-th tuple}},\dots,\overbrace{\rho_{n,\sigma_n^{-1}(a_n)}(1),\rho_{n,\sigma_n^{-1}(a_n)}(2),\dots,\rho_{n,\sigma_n^{-1}(a_n)}(k_{n,\sigma_n^{-1}(a_n)})}^{(a_1+\cdots+a_{\tau(1)-1}+a_{\tau(1)})\text{-th tuple}}}_{\tau(n)\text{-th } n\text{-block}})$$

The above tuple can be (again, horribly) understood as

$$(\dots,\underbrace{\rho_{\tau^{-1}(i),\sigma^{-1}_{\tau^{-1}(i)}(j)}(1),\rho_{\tau^{-1}(i),\sigma^{-1}_{\tau^{-1}(i)}(j)}(2),\dots,\rho_{\tau^{-1}(i),\sigma^{-1}_{\tau^{-1}(i)}(j)}(k_{\tau^{-1}(i),\sigma^{-1}_{\tau^{-1}(i)}(j)}),\dots}_{(a_1+\cdots+a_{i-1}+j)\text{-th block}})$$

Which shows that

$$(\tau\circ_{a_1,\dots,a_n}(\sigma_1,\sigma_2,\dots,\sigma_n))\circ_{k_{1,1},\dots,k_{1,a_1},\dots,k_{n,1},\dots,k_{n,a_n}}\circ(\rho_{1,1},\dots,\rho_{n,a_n})$$
$$=$$
$$\tau\circ_{(k_{1,1}+\cdots+k_{1,a_1}),\dots,(h_{n,1}+\cdots+k_{n,a_n})}(\sigma_1\circ_{k_{1,1},\dots,k_{1,a_1}}(\rho_{1,1},\dots,\rho_{1,a_1}),\dots,\sigma_n\circ_{k_{n,1},\dots,k_{n,a_n}}(\rho_{n,1},\dots,\rho_{n,a_n}))$$

so that **NS-OP1** is satisfied. Now verifying **NS-OP2** is simple; note that as $S_1$ has one element, we are forced to identify our identity element as $\sigma_1$, the unique permutation of one element that doesn't do anything. Then for any $\tau\in S_n$, we of course have that $\tau\circ_{1,1,\dots,1}(\sigma_1,\sigma_1,\dots\sigma_1)=\tau$, as each element is unchanged by $\sigma_1$ before $\tau$ is applied. We also know that $\sigma_1\circ_n(\tau)=\tau$, since this is just applying $\tau$ and then applying the trivial block permutation to the $n$ elements.

Now we show **S-OP1**. As we need a right action of $S_n$ on the $n$-th level of our operad, which also happens to be $S_n$, an evident choice would be to just take the group product. Hence for any $\sigma \in S_n$, we say $\tau \in S_n$ acts on $\sigma$ to give rise to

$$(\sigma \cdot \tau) = \sigma \circ \tau$$

which is clearly in $S_n$.

To demonstrate **S-OP1**, let $\tau, \rho \in S_n$, and $\sigma_1 \in S_{a_1}, \dots, \sigma_n \in S_{a_n}$ for $a_i \in \mathbb{N}$. To compute $(\tau \cdot \rho) \circ_{a_1,\dots,a_n} (\sigma_1, \dots, \sigma_n)$, denote an (ordered) tuple of the first $a_1 + \cdots + a_n$ integers as

$$({\color{red}1, 2, \dots, a_1}, \dots, {\color{blue}1', 2', \dots, a_n}).$$

Then we see that $(\tau \cdot \rho) \circ_{a_1,\dots,a_n} (\sigma_1, \dots, \sigma_n)$ acts on the tuple to give rise to

$$(\sigma'_{\rho^{-1}(\tau^{-1}(1))}(1), \dots, \sigma'_{\rho^{-1}(\tau^{-1}(1))}(a_{\rho^{-1}(\tau^{-1}(1))}), \dots, \sigma'_{\rho^{-1}(\tau^{-1}(n))}(1), \dots, \sigma'_{\rho^{-1}(\tau^{-1}(n))}(a_{\rho^{-1}(\tau^{-1}(n))}))$$

On the other hand we need to also compute $(\tau \circ_{a_{\rho^{-1}(1)},\dots,a_{\rho^{-1}(n)}} (\sigma_{\rho^{-1}(1)}, \dots, \sigma_{\rho^{-1}(n)})) \cdot \rho'$ where $\rho'$ is the evident block permutation. However, this is really just $(\tau \circ_{a_{\rho^{-1}(1)},\dots,a_{\rho^{-1}(n)}} (\sigma_{\rho^{-1}(1)}, \dots, \sigma_{\rho^{-1}(n)})) \circ \rho'$; below we see that its action on an ordered $a_1 + \cdots + a_n$ tuple is as we would expect.

$$\begin{gathered}
({\color{red}1, 2, \dots, a_1}, \dots, {\color{blue}1', 2', \dots, a_n}) \\
\downarrow{\scriptstyle \rho'} \\
(1', 2', \dots, a_{\rho^{-1}(1)}, \dots, 1', 2', \dots, a_{\rho^{-1}(n)}) \\
\downarrow{\scriptstyle (\tau \circ_{a_{\rho^{-1}(1)},\dots,a_{\rho^{-1}(n)}} (\sigma_{\rho^{-1}(1)},\dots,\sigma_{\rho^{-1}(n)}))} \\
(\sigma'_{\rho^{-1}(\tau^{-1}(1))}(1), \dots, \sigma'_{\rho^{-1}(\tau^{-1}(1))}(a_{\rho^{-1}(\tau^{-1}(1))}), \dots, \sigma'_{\rho^{-1}(\tau^{-1}(n))}(1), \dots, \sigma'_{\rho^{-1}(\tau^{-1}(n))}(a_{\rho^{-1}(\tau^{-1}(n))}))
\end{gathered}$$

Therefore we see that

$$(\tau \cdot \rho) \circ_{a_1,\dots,a_n} (\sigma_1, \dots, \sigma_n) = (\tau \circ_{a_{\rho^{-1}(1)},\dots,a_{\rho^{-1}(n)}} (\sigma_{\rho^{-1}(1)}, \dots, \sigma_{\rho^{-1}(n)})) \cdot \rho'$$

so that **S-OP1** is satisfied. We just now need to show **S-OP2** is satisfied, which is nearly immediate. We will however not pretend we're too good to show this and demonstrate it anyways. For each $\sigma_i \in S_{a_i}$, pick $\rho_i \in S_{a_i}$. Observe that $\tau \circ_{a_1,\dots,a_n} (\sigma_1 \cdot \rho_1, \dots, \sigma_n \cdot \rho_n)$

$$\begin{gathered}
({\color{red}1, 2, \dots, a_1}, \dots, {\color{blue}1', 2', \dots, a_n}) \\
\downarrow{\scriptstyle (\tau\cdot\rho) \circ_{a_1,\dots,a_n} (\sigma_1,\dots,\sigma_n)} \\
(\overbrace{\sigma_{\tau^{-1}(1)}(\rho_{\tau^{-1}(1)}(1)), \sigma_{\tau^{-1}(1)}(\rho_{\tau^{-1}(1)}(2)), \dots, \sigma_{\tau^{-1}(1)}(\rho_{\tau^{-1}(1)}(a_{\tau^{-1}(1)}))}^{\text{1st block}}, \dots \\
\dots, \overbrace{\sigma_{\tau^{-1}(n)}(\rho_{\tau^{-1}(n)}(1)), \sigma_{\tau^{-1}(n)}(\rho_{\tau^{-1}(n)}(2)), \dots, \sigma_{\tau^{-1}(n)}(\rho_{\tau^{-1}(n)}(a_{\tau^{-1}(n)}))}^{n\text{-th block}})
\end{gathered}$$

returns the same result as $(\tau \circ_{a_1,\dots,a_n} (\sigma_1, \dots, \sigma_n)) \cdot (\rho_1, \dots, \rho_n)$

$$
\begin{gathered}
(1, 2, \dots, a_1, \dots, 1', 2', \dots, a_n) \\
\Big\downarrow {\scriptstyle (\rho_1, \dots, \rho_n)} \\
(\rho_1^{-1}(1), \rho_1^{-1}(2), \dots, \rho_1^{-1}(a_1), \dots, \rho_n^{-1}(1), \rho_n^{-1}(2), \dots, \rho_n^{-1}(a_n)) \\
\Big\downarrow {\scriptstyle \tau \circ_{a_1, \dots, a_n} (\sigma_1, \dots, \sigma_n)} \\
(\overbrace{\sigma_{\tau^{-1}(1)}(\rho_{\tau^{-1}(1)}(1)), \sigma_{\tau^{-1}(1)}(\rho_{\tau^{-1}(1)}(2)), \dots, \sigma_{\tau^{-1}(1)}(\rho_{\tau^{-1}(1)}(a_{\tau^{-1}(1)}))}^{\text{1st block}}, \dots \\
\dots, \overbrace{\sigma_{\tau^{-1}(n)}(\rho_{\tau^{-1}(n)}(1)), \sigma_{\tau^{-1}(n)}(\rho_{\tau^{-1}(n)}(2)), \dots, \sigma_{\tau^{-1}(n)}(\rho_{\tau^{-1}(n)}(a_{\tau^{-1}(n)}))}^{n\text{-th block}})
\end{gathered}
$$

since $(\tau \circ_{a_1, \dots, a_n} (\sigma_1, \dots, \sigma_n)) \cdot (\rho_1, \dots, \rho_n) = (\tau \circ_{a_1, \dots, a_n} (\sigma_1, \dots, \sigma_n)) \circ (\rho_1, \dots, \rho_n)$ in our case. As we have that **S-OP2** is satisfied, we have that $\text{Assoc}_n = S_n$ is a symmetric operad.

---

**Definition 9.1.5.** An **morphism of operads** $F : X \longrightarrow Y$ between two (symmetric) operads $X, Y$ with units $I \in X_1$ and $J \in Y_1$ and $S_n$ group actions $\cdot$ and $*$ is a family of maps $F_n : X_n \longrightarrow Y_n$ such that

**(M-OP1)** $F_1(I) = J$

**(M-OP2)** If $f \in X_n$ and $g_1 \in X_{a_1}, \dots, g_n \in X_{a_n}$ for $a_i \in \mathbb{N}$, then

$$F_{a_1 + \cdots + a_n}(f \circ_{a_1, \dots, a_n} (g_1, \dots, g_n)) = F_n(f) \circ_{a_1, \dots, a_n} (F_{a_1}(g_1), \dots, F_{a_n}(g_n))$$

**(M-OP3)** If $f \in X_n$ and $\tau \in S_n$, then

$$F_n(f \cdot \tau) = F_n(f) * \tau$$

Note: in the case where $X, Y$ are symmetric operads, we define a morphism between $X$ and $Y$ to be a family of maps $F_n : X_n \longrightarrow Y_n$ such that only **M-OP1** and **M-OP2** hold.

**Definition 9.1.6.** A **algebra over an Operad** $X$ is a morphism of operads $F : X \longrightarrow \text{End}_A$ where $A$ is some set. Spelled out, this is a mapping

$$
\begin{aligned}
F_n : X_n &\longrightarrow \text{Hom}_{\textbf{Set}}(A^n, A) \\
f &\mapsto F_n(f) : A^n \longrightarrow A
\end{aligned}
$$

so that we're mapping elements of our operad to $n$-ary operations over $A$. This mapping also requires that

1. $F_1(I) = \text{id}_A : A \longrightarrow A$
2. For $f \in X_n$, $g_i \in X_{a_i}$ for $i = 1, 2, \dots, n$,

$$F_{a_1 + \cdots + a_n}(f \circ_{a_1, \dots, a_n} (g_1, \dots, g_n)) = F_n(f) \circ'_{a_1, \dots, a_n} (F_{a_1}(g_1), \dots, F_{a_n}(g_n)).$$

Diagrammatically, this means the following diagrams commutes:

$$\begin{tikzcd}
X_n \times (X_{a_1} \times \cdots \times X_{a_n}) \arrow[r, "(F_n, F_{a_1}, \dots, F_{a_n})"] \arrow[d, "\circ_{a_1,\dots,a_n}"'] & \mathrm{Hom}(A^n, A) \times (\mathrm{Hom}(A^{a_1}, A) \times \cdots \times \mathrm{Hom}(A^{a_n}, A)) \arrow[d, "\circ'_{a_1,\dots,a_n}"] \\
X_{a_1+\cdots+a_n} \arrow[r, "F_{a_1+\cdots+a_n}"'] & \mathrm{Hom}(A^{a_1+\cdots+a_n}, A)
\end{tikzcd}$$

Or, more visually,

$$\begin{tikzcd}
A^{a_1} \times A^{a_2} \times \cdots \times A^{a_n} \arrow[d, "F(f \circ_{a_1,\dots,a_n}(g_1,\dots,g_n))"'] \\
A
\end{tikzcd}
=
\begin{tikzcd}
A^{a_1} \arrow[d, "F_{a_1}(g_1)"'] & A^{a_2} \arrow[d] & \cdots & A^{a_n} \arrow[d, "F_{a_n}(g_n)"] \\
A \times & A \times \arrow[d, "F(f)"] & \overbrace{\cdots}^{n \text{ times}} & \times A \\
& A & &
\end{tikzcd}$$

3. Finally, we have that if $\tau \in S_n$, then for $f \in X_n$ and $(a_1, \dots, a_n) \in A^n$, then

$$F_n(f \cdot \tau)(a_1, \dots, a_n) = (F_n(f) * \tau)(a_1, \dots, a_n) = F_n(f)(a_{\tau(1)}, \dots, a_{\tau(n)}).$$

**Definition 9.1.7.** Let $X$ be an operad. A **morphism** $\Phi : F \longrightarrow G$ **between algebras** $F : X \longrightarrow \mathrm{End}_A$ and $G : X \longrightarrow \mathrm{End}_B$ over $X$ is a function $\varphi : A \longrightarrow B$ such that, for $f \in X_n$ and $(a_1, \dots, a_n) \in A^n$,

$$\varphi(F_n(f)(a_1, \dots, a_n)) = G(f)(\varphi(a_1), \dots, \varphi(a_n))$$

The above relation can be more conveniently expressed as the diagram below commuting:

$$\begin{tikzcd}
A^n \arrow[r, "F_n(f)"] \arrow[d, "(\varphi,\varphi,\dots,\varphi)"'] & A \arrow[d, "\varphi"] \\
B^n \arrow[r, "G_n(f)"'] & B
\end{tikzcd}$$

which must hold for all $f \in X_n$ with $n \in \mathbb{N}$. Now suppose that for an operad $X$ we have three algebras

$$F : X \longrightarrow \mathrm{End}_A \quad G : X \longrightarrow \mathrm{End}_B \quad H : X \longrightarrow \mathrm{End}_C$$

such that $\Phi : F \longrightarrow G$ and $\Psi : G \longrightarrow H$ are morphisms of algebras given by functions $\varphi : A \longrightarrow B$ and $\psi : B \longrightarrow C$. A natural question is whether or not one can define a morphism $\Psi \circ \Phi : F \longrightarrow G$. This is however immediate upon realization that we can stack the diagrams to see that $\Phi \circ \Psi : F \longrightarrow H$ is a morphism of algebras.

$$\begin{tikzcd}
A^n \arrow[r, "F_n(f)"] \arrow[d, "(\varphi,\varphi,\dots,\varphi)"'] & A \arrow[d, "\varphi"] \\
B^n \arrow[r, "G_n(f)"'] \arrow[d, "(\psi,\psi,\dots,\psi)"'] & B \arrow[d, "\psi"] \\
C^n \arrow[r, "H_n(G_n(f))"'] & C
\end{tikzcd}$$

As a result, if we are given an operad $X$, we can create a category $\mathbf{Alg}_X$ whose objects are algebras $\Phi : X \longrightarrow \text{End}_A$ and whose morphisms are morphisms between such algebras. These categories actually return ordinary categories that we've dealt with in the past.

---

**Example 9.1.8.** Consider the operad $\text{Assoc}_n = S_n$. Then we have that

$$\mathbf{Alg}_{\text{Assoc}_n} \cong \mathbf{Mon}$$

where **Mon** is the category of monoids. (In terms of set theory, we're being sloppy; but if anyone challenges this we can just pull out a Grothendieck universe and satisfy their demands.) To demonstrate this isomorphism we must produce a pair of inverse functors between these categories.

Before we do that, first consider an object in this category, which is a family of functions $F_n : S_n \longrightarrow \text{Hom}_{\mathbf{Set}}(A^n, A)$ for some set $A$. To save some space, denote $\text{Hom}_{\mathbf{Set}}(A^n, A)$ as $[A^n, A]$. Then the fact that $F : \mathbf{Assoc}_n \longrightarrow \text{End}_A$ is an algebra gives us that the diagram on the left commutes.

$$\begin{tikzcd}
S_2 \times (S_2 \times S_1) \arrow[r] \arrow[d] & {[A^2, A] \times \Big([A^2, A], \times [A, A]\Big)} \arrow[d] \\
S_3 \arrow[r] & \text{Hom}(A^3, A)
\end{tikzcd}
\qquad
\begin{tikzcd}
(e_2, e_2, e_1) \arrow[r, mapsto] \arrow[d, mapsto] & (\mu_2, \mu_2, \text{id}_A) \arrow[d, mapsto] \\
e_3 \arrow[r, mapsto] & \mu_3 = \mu_2(\mu_2, \text{id}_A)
\end{tikzcd}$$

As this diagram commutes, we can follow the specific path which is taken by the identity elements $e_2 \in S_2$ and $e_1 \in S_1$. If we denote $F_n(e_n) = \mu_n : A^n \longrightarrow A$, then we see that $\mu_3 = \mu_2(\mu_2, \text{id}_A)$. Note that in particular, $\mu_1 = \text{id}_A$ by hypothesis. Hence for $a, b, c \in A$, we see that $\mu_3 = \mu_2(\mu_2(a, b), c)$. Conversely, we can repeat the same thing with $S_1$ and $S_2$ swapped, and obtain a commutative diagram on the left:

$$\begin{tikzcd}
S_2 \times (S_1 \times S_2) \arrow[r] \arrow[d] & {[A^2, A] \times \Big([A, A], \times [A^2, A]\Big)} \arrow[d] \\
S_3 \arrow[r] & \text{Hom}(A^3, A)
\end{tikzcd}
\qquad
\begin{tikzcd}
(e_2, e_1, e_2) \arrow[r, mapsto] \arrow[d, mapsto] & (\mu_2, \text{id}_A, \mu_2) \arrow[d, mapsto] \\
e_3 \arrow[r, mapsto] & \mu_3 = \mu_2(\text{id}_A, \mu_2)
\end{tikzcd}$$

and following the identity elements again grants us that $\mu_3 = \mu_2(\text{id}_A, \mu_2)$. Hence we see that for $a, b, c \in A$ $\mu_3(a, b, c) = \mu_2(a, \mu_2(b, c))$. All together we have that

$$\mu_2(\mu_2(a, b), c) = \mu_2(a, \mu_2(b, c)).$$

What does this mean? Perhaps this will make it more clear: denote $\mu_2(a, b) = a \cdot b$. Then this means that

$$(a \cdot b) \cdot c = a \cdot (b \cdot c).$$

This means that we've proved that $A$ is a set equipped with a binary operator $\mu_2 : A \times A \longrightarrow A$ which is associative! This is almost a monoid; we're just missing an identity element. However, note that

# 9.2 General Operads in Symmetric Monoidal Categories

Every time we find ourselves working in **Set**, we should feel a great deal of shame and embarrassment. Before anyone catches us, we can atone for our sins by drawing diagrams that avoid specific reference to the element of the sets, thereby transitioning our work to an arbitrary category. Given our previous work, we can do this; but what were the main ingredients? Note that we basically only needed the properties of **Set** and its cartesian product. Given this, and the fact that **Set** is symmetric monoidal given the cartesian product, we can largely generalize our previous work to arbitrary symmetric monoidal categories.

**Definition 9.2.1.** Let $(\mathcal{C}, \otimes, I)$ be a symmetric monoidal category. A (symmetric) **operad** $X$ over $\mathcal{C}$ is a family of objects $\{X_n\}_{n\in\mathbb{N}}$, in $\mathcal{C}$, where each $X_n$ has a group action by $S_n$ and with

1. A unit morphism $\eta : I \longrightarrow X_1$
2. For each $n \in \mathbb{N}$ and $a_i \in \mathbb{N}$ where $i = 1, 2, \dots, n$, a composition morphism

$$\mu : X_n \otimes X_{a_1} \otimes \cdots \otimes X_{a_n} \longrightarrow X_{a_1+\cdots+a_n}$$

subject to the associativity, identity, and equivariance axioms outlined below.

**(OP1) Associativity.** Let $n \geq 0$ and choose $a_i \geq 0$ for $i = 1, 2, \dots, n$. Further, for each $a_i$, choose $k_{i,j} \geq 0$ for $j = 1, 2, \dots, a_i$. Let $\gamma$ be the isomorphism which rearranges the factors of the tensor product as below:

$$\begin{gathered}\gamma : (X_n \otimes X_{a_1} \otimes \cdots \otimes X_{a_n}) \otimes X_{k_{1,1}} \otimes \cdots \otimes X_{k_{1,a_1}} \otimes \cdots \otimes X_{k_{n,1}} \otimes \cdots \otimes X_{k_{n,a_n}} \\ \xrightarrow{\sim} \\ X_n \otimes (X_{a_1} \otimes X_{k_{1,1}} \otimes \cdots \otimes X_{k_{1,a_1}}) \otimes \cdots \otimes (X_{a_n} \otimes X_{k_{n,1}} \otimes \cdots \otimes X_{k_{n,a_n}})\end{gathered}$$

Then we demand that the diagram below commutes.

$$\begin{tikzcd}
\left(X_n \otimes \bigotimes_{i=1}^{n} X_{a_i}\right) \otimes \left(\bigotimes_{i=1}^{n}\bigotimes_{j=1}^{a_i} X_{k_{i,j}}\right) \arrow[rr, "\gamma"] \arrow[d, "\mu_{a_1,\dots,a_n}\otimes 1"'] & & X_n \otimes \bigotimes_{i=1}^{n}\left(X_{a_i} \bigotimes_{j=1}^{a_i} X_{k_{i,j}}\right) \arrow[d, "1_{X_n}\otimes\mu\otimes\cdots\otimes\mu"] \\
X_{a_1+\cdots+a_n} \otimes \left(\bigotimes_{i=1}^{n}\bigotimes_{j=1}^{a_i} X_{k_{i,j}}\right) \arrow[dr, "\mu_{(k_{1,1}+\cdots+k_{1,a_1}),\dots,(k_{n,1}+\cdots+k_{n,a_n})}"'] & & X_n \otimes X_{k_{1,1}+\cdots+k_{1,a_1}} \otimes \cdots \otimes X_{k_{n,1}+\cdots+k_{n,a_n}} \arrow[dl, "\mu_{(k_{1,1}+\cdots+k_{1,a_1})+\cdots+(k_{n,1}+\cdots+k_{n,a_n})}"] \\
& X_{k_{1,1}+\cdots+k_{1,a_1}+\cdots+k_{n,1}+\cdots+k_{n,a_n}} &
\end{tikzcd}$$

**(OP2) Identity.** Letting $A$ be an arbitrary object of $\mathcal{C}$, let $\lambda : I \otimes A \xrightarrow{\sim} A$ and $\rho : A \otimes I \xrightarrow{\sim} A$

as the left and right unitors in our symmetric monoidal category. Then the diagrams below must hold for all $n \ge 0$.

$$
\begin{tikzcd}
I \otimes X_n \arrow[r, "\eta \otimes 1_{X_n}"] \arrow[dr, "\lambda"'] & X_1 \otimes X_n \arrow[d, "\mu_n"] \\
& X_n
\end{tikzcd}
\qquad
\begin{tikzcd}
X_n \otimes I^{\otimes n} \arrow[r, "1_{X_n} \otimes \eta^{\otimes n}"] \arrow[dr, "\rho^{\otimes n}"'] & X_n \otimes X_1^{\otimes n} \arrow[d, "\mu"] \\
& X_n
\end{tikzcd}
$$

**(OP3) Equivariance 1.** Let $\tau \in S_n$, and let $\tau^*$ be the isomorphism $\tau^* : X_{a_1} \otimes \cdots \otimes X_{a_n} \xrightarrow{\sim} X_{\tau(a_1)} \otimes \cdots \otimes X_{\tau(a_n)}$ and by abuse of notation denote $\tau$ as the morphism $\tau : X_n \longrightarrow X_n$ which is given by the group action. Then the diagram below must commute.

$$
\begin{tikzcd}
X_n \otimes X_{a_1} \otimes \cdots \otimes X_{a_n} \arrow[rr, "\tau \otimes 1_{X_{a_1}} \otimes \cdots \otimes 1_{X_{a_n}}"] \arrow[d, "1_{X_n} \otimes \tau^*"'] & & X_n \otimes X_{a_1} \otimes \cdots \otimes X_{a_n} \arrow[d, "\mu_{a_1, \dots, a_n}"] \\
X_n \otimes X_{\tau(a_1)} \otimes \cdots \otimes X_{\tau(a_n)} \arrow[dr, "\mu_{\tau(a_1), \dots, \tau(a_n)}"'] & & X_{a_1 + \cdots + a_n} \\
& X_{\tau(a_1) + \cdots + \tau(a_n)} \arrow[ur, "\tau'"'] &
\end{tikzcd}
$$

Here, $\tau'$ is the block permutation described below:

$$
(\overbrace{1, 2, \dots, a_1}^{\text{1st block}}, \dots, \overbrace{a_1 + \cdots + a_i + 1, \dots a_1 + \cdots + a_{i+1}}^{i\text{-th block}}, \dots \overbrace{a_1 + \cdots + a_{n-1} + 1, \dots, a_1 + \cdots + a_n}^{n\text{-th block}})
$$

$$\mapsto$$

$$
(\dots, \overbrace{1, 2, \dots, a_1}^{\tau(1)\text{-th block}}, \dots, \overbrace{a_1 + \cdots + a_i + 1, \dots, a_1 + \cdots + a_{i+1}}^{\tau(i)\text{-th block}}, \dots, \overbrace{a_1 + \cdots + a_{n-1} + 1, \dots, a_1 + \cdots + a_n}^{\tau(n)\text{-th block}}, \dots
$$

**(OP4) Equivariance 2.** Let $\sigma_i \in S_{a_i}$ for $i = 1, 2, \dots, n$. By abuse of notation, denote $\sigma_i : X_{a_i} \longrightarrow X_{a_i}$ to be the map given by the group action. Then we have that

$$
\begin{tikzcd}
X_n \otimes X_{a_1} \otimes \cdots \otimes X_{a_n} \arrow[r, "1_{X_n} \otimes \sigma_1 \otimes \cdots \otimes \sigma_n"] \arrow[d, "\mu_{a_1, \dots, a_n}"'] & X_n \otimes X_{a_1} \otimes \cdots \otimes X_{a_n} \arrow[d, "\mu_{a_1, \dots, a_n}"] \\
X_{a_1 + \cdots + a_n} \arrow[r, "(\sigma_1, \sigma_2, \dots, \sigma_n)"'] & X_{a_1 + \cdots + a_n}
\end{tikzcd}
$$

where $(\sigma_1, \sigma_2, \ldots, \sigma_n)$ is the permutation in $S_{a_1+\cdots+a_n}$ defined as below.

$$(\overbrace{1, 2, \ldots, a_1}^{\text{1st block}}, \ldots, \overbrace{a_1 + \cdots + a_{n-1} + 1, , \ldots, a_1 + \cdots + a_{n-1} a_n}^{n\text{-th block}})$$

$$\mapsto$$

$$(\underbrace{\sigma_1(1), \sigma_1(2), \ldots, \sigma_1(a_1)}_{\text{1st block}}, \ldots, \underbrace{a_1 + \cdots + a_{n-1} + \sigma_n(1), \ldots, a_1 + \cdots + a_{n-1} + \sigma_n(a_n)}_{n\text{-th block}})$$

---

**Example 9.2.2.** As before, we can create an endomorphism operad. That is, if we let $\mathcal{C}$ be a symmetric monoidal category, then we can let $\text{End}_A(n) = \text{Hom}_{\mathcal{C}}(A^{\otimes n}, A)$. Then $u : I \longrightarrow \text{Hom}_{\mathcal{C}}(X, X)$ is defined to be the unique map to the identity. Given $f \in \text{End}_A(n)$ and $g_i \in \text{End}_A(a_i)$ where $a_i \in \mathbb{N}$ for $i = 1, 2, \ldots, n$, then we define our composition pointwise:

$$f \circ_{a_1,\ldots,a_n} (g_1, \ldots, g_n) = f \circ (g_1 \otimes \cdots \otimes g_n).$$

Finally, given $\sigma \in S_n$, we can define a group action by assigning $f \cdot \sigma$ to the morphism which rearranges the positioning of $A^{\otimes n}$ according to the permutation $\sigma$. With these hypotheses one can check that the axioms of an operad are satisfied as we did in the previous section when $\mathcal{C} = \mathbf{Set}$.

---

## 9.3 Partial Composition: Restructuring Operads

After one stares at the definition of an operad for quite some time, they will realize that the vast and mysterious diagrams and indices are really just for booking keeping, and that the idea is actually rather quite intuitive. And of this bookkeeping is what makes operads a bit annoying; we are constantly having to think about an arbitrarily long tensor products. However, Freese has pointed out in his text that we can actually rephrase the language of operads more simply by replacing the arbitrarily long composition morphism with a *partial composition morphism*. However, this itself is not trivial.

Let $X$ be a set, and consider the endomorphism operad $\text{End}_X(n)$. For any $f \in \text{Hom}_{\textbf{Set}}(X^n, X)$, we can choose $g_i \in \text{Hom}_{\textbf{Set}}(X^{a_i}, X)$ for $a_i \in \mathbb{N}$ with $i = 1, 2, \dots, n$. Composition can then be defined pointwise:

$$\begin{aligned} f \circ_{a_1,\dots,a_n} (g_1, \dots, g_n)(x_1, \dots, x_{a_1}, \dots, x_{a_1+\cdots+a_{n-1}+1}, \dots, x_{a_1+\cdots+a_n}) \\ = f(g_1(x_1, \dots x_{a_1}), \dots, g_n(x_{a_1+\cdots+a_n+1}, \dots, x_{a_1+\cdots+a_n})) \end{aligned}$$

However, what if we decided to build this function another way; perhaps, handling one $g_i$ at a time? The way we could do this is by inserting a $g_i$ one at a time:

$$(f, g_i) \mapsto f(\underbrace{x_1, \dots, x_{k-1}}_{k-1}, \overbrace{g_i(x'_1, \dots, x'_{a_i})}^{k\text{-th spot}}, \underbrace{x_{k+1}, \dots, x_n}_{n-(k+1)})$$

Given that we'd have a total of $(n + a_i - 1)$-many inputs, this then defines a composition operator

$$\circ_k : X^n \times X^{a_i} \longrightarrow X^{n+a_i-1}$$

for each $n, a_i \geq 0$. We can then repeatedly apply this composition operator to build the same function that our operadic composition does.

**Definition 9.3.1.** Let $X$ be an operad in a symmetric monoidal category $\mathcal{C}$. Then for each $n, m \geq 0$, we define the **partial composition operator** $\circ_k : X_n \otimes X_m \longrightarrow X_{n+m-1}$ as the composition of the arrows pictured below.

$$\begin{tikzcd} X_m \otimes X_n \arrow[r, "\sim"] \arrow[d, dashed, "\circ_k"'] & X_m \otimes (I \otimes \cdots \otimes \overset{k\text{-th factor}}{X_n} \otimes \cdots \otimes I) \arrow[d, "1_{X_m} \otimes \eta \otimes \cdots \otimes 1_{X_n} \otimes \cdots \otimes \eta"] \\ X_{n+m-1} & X_m \otimes X_1 \otimes \cdots \otimes X_n \otimes \cdots \otimes X_1 \arrow[l, "\mu_{1,\dots,1,n,1\dots,1}"] \end{tikzcd}$$

In other words, the partial composition operator $\circ_k$ on $X_m \otimes X_n$ is the same as our original composition operator $\mu$ applied to $X_m \otimes X_1 \otimes \cdots \otimes X_n \otimes \cdots \otimes X_1$.

It was Fresse who demonstrated in his gigantic text that the partial composition operator can equivalently construct operads. The strategy he used is as follows: we first investigate what properties (i.e. diagrams) that the partial composition operator satisfies. Then, we forget that we ever had on operad, but we rather consider a sequence of objects which are basically operads, but whose composition operator has now been replaced by the partial composition operator. Fresse showed that these objects then form a category, and that this category is isomorphic to the category of operads, thereby demonstrating an equivalence of operad definitions and paving the way for simpler calculations in demonstrating that something is an operad.

Thus we demonstrate properties of the partial composition operator. Let $X$ be an operad and recall the associativity pentagon given in **OP1**. In the associativity diagram, replace $X_{a_i} = X_1$ except $X_{a_p} = X_r$ for some $p \leq n$, and set $X_{k_{i,j}} = X_1$ except for $X_{k_{p,q}} = X_s$ for some $q \leq a_p$. Then we get the commutative diagram below.

$$
\begin{tikzcd}[column sep=small]
X_n \otimes (\underbrace{X_1 \otimes \cdots \otimes \overset{k\text{-th spot}}{X_r} \otimes \cdots \otimes X_1}_{n \text{ factors}}) \otimes (\underbrace{X_1 \otimes \cdots \otimes \overset{\ell\text{-th spot}}{X_s} \otimes \cdots \otimes X_1}_{r \text{ factor}}) \arrow[r, "\sim"] \arrow[d, "\mu_{1,\dots,r,\dots,1} \otimes (1_{X_1} \otimes \cdots \otimes 1_{X_r} \otimes \cdots \otimes 1_{X_1})"'] & X_n \otimes (\underbrace{X_1 \otimes \cdots \otimes \overset{k\text{-th spot}}{(X_r \otimes (X_1 \otimes \cdots \otimes X_s \otimes \cdots \otimes X_1))} \otimes \cdots \otimes X_1}_{n \text{ factors}}) \arrow[d, "1_{X_n} \otimes 1_{X_1} \otimes \cdots \otimes \circ_\ell \otimes \cdots \otimes 1_{X_1}"] \\
X_{n+r-1} \otimes (X_1 \otimes \cdots \otimes X_s \otimes \cdots \otimes X_1) \arrow[dr, "\mu_{1,1,\dots,s,\dots,1}"'] & X_n \otimes (X_1 \otimes \cdots \otimes X_{r+s-1} \otimes \cdots \otimes X_1) \arrow[d, "\mu_{1,1,\dots,r+s-1,\dots,1}"] \\
& X_{n+r+s-2}
\end{tikzcd}
$$

With similar substitutions, we also get that the diagram below commutes.

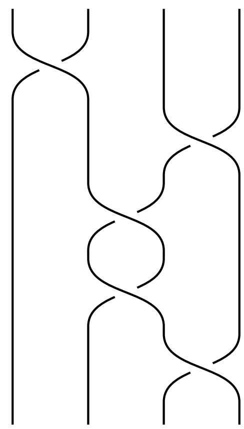
$$
\begin{tikzcd}[column sep=small]
X_n \otimes (\underbrace{X_1 \otimes \cdots \otimes \overset{k\text{-th spot}}{X_r} \otimes \cdots \otimes X_1}_{n \text{ factors}} \otimes \underbrace{X_1 \otimes \cdots \otimes \overset{\ell\text{-th spot}}{X_s} \otimes \cdots \otimes X_1}_{(n+r-1) \text{ factors}}) \arrow[r, "\sim"] \arrow[d, "\mu_{1,1,\dots,r,\dots,1}"'] & X_n \otimes (\underbrace{X_1 \otimes \cdots \otimes \overset{k\text{-th spot}}{X_s} \otimes \cdots \otimes X_1}_{n \text{ factors}} \otimes \underbrace{X_1 \otimes \cdots \otimes \overset{\ell\text{-th spot}}{X_r} \otimes \cdots \otimes X_1}_{(n+r-1) \text{ factors}}) \arrow[d, "\mu_{1,1,\dots,s,\dots,1}"] \\
X_{n+r-1} \otimes (I \otimes \cdots \otimes X_r \otimes \cdots \otimes I) \arrow[dr, "\mu_{1,1,\dots,s,\dots,1}"'] & X_{n+s-1} \otimes (I \otimes \cdots \otimes X_r \otimes \cdots \otimes I) \arrow[d, "\mu_{1,1,\dots,r,\dots,1}"] \\
& X_{m+n+r-2}
\end{tikzcd}
$$

## 9.4 The Braid Groups Form a (nonsymmetric) Operad

Recall that the $n$-th braid group $B_n$ is the collection of all possible braidings of $n$-strands, forming a group under composition. Each braid group has the presentation

$$B_n = \left\langle \sigma_1, \dots, \sigma_{n-1} \mid \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}^{(1)}, \sigma_i\sigma_j = \sigma_j\sigma_i^{(2)} \right\rangle$$

where (1) holds only when $1 \leq i \leq n-2$ and (2) hold only when $|i-j| > 1$. Below is the braid $\sigma_1\sigma_3\sigma_2\sigma_2\sigma_3$, where we envision application of the generators starting from the left and going to the right.



Each braid group has a natural projection mapping $\pi : B_n \longrightarrow S_n$, where each braid is sent to the underlying permutation. The kernel of this map is the pure braid group, which doesn't change the permutation. However, recall that $S_n$ is a symmetric operad, whose composition is given by a block permutation. That is, given a permutation $S_n$, and $n$-many other permutations $\sigma_1 \in S_{a_1}, \dots, \sigma_n \in S_{a_n}$, we can form a permutation in $S_{a_1+\cdots+a_n}$.

$$(1, 2, \dots, a_1, 1', 2', \dots, a'_2, \dots, 1', 2', \dots, a'_n)$$

$$\downarrow \sigma_1 \qquad \downarrow \sigma_2 \qquad \cdots \qquad \downarrow \sigma_n$$

$$(\underbrace{\sigma_1(1), \sigma_1(2)), \dots \sigma_1(a_1)}_{\text{1st block}}, \underbrace{\sigma'_1(1), \sigma'_1(2)), \dots \sigma'_1(a_1)}_{\text{2nd block}}, \dots, \underbrace{\sigma'_n(1), \sigma'_n(2) \dots, \sigma'_n(a_n)}_{a_n\text{-th block}})$$

$$\downarrow \tau$$

$$(\dots, \underbrace{\sigma_1(1), \sigma_1(2)), \dots \sigma_1(a_1)}_{\tau(1)\text{-th block}}, \dots, \underbrace{\sigma'_1(1), \sigma'_1(2)), \dots \sigma'_1(a_1)}_{\tau(2)\text{-th block}}, \dots, \underbrace{\sigma'_n(1), \sigma'_n(2), \dots, \sigma'_n(a_n)}_{\tau(a_n)\text{-th block}}, \dots, )$$

This then suggests the idea that there exists an operadic composition for braids; and such an observation checks out. Given a braid $\beta \in B_n$, and $n$-many other braids $\alpha_1 \in B_{a_i}, \dots, \alpha_n \in B_{a_n}$, we can form a braid in $B_{a_1+\cdots+a_n}$. The operadic composition is analogous to what we had before with permutations; we're going to stick braids inside of braids.

**Definition 9.4.1.** (**Topological.**) Let $\beta \in B_n$ be a braid. We say that the $i, (i+1), \dots (i+k)$-th strands form a **cable** if there exist a cylinder (depends on ambient space; need to decide one for consistency) which is disjoint from all other strands of $\beta$.

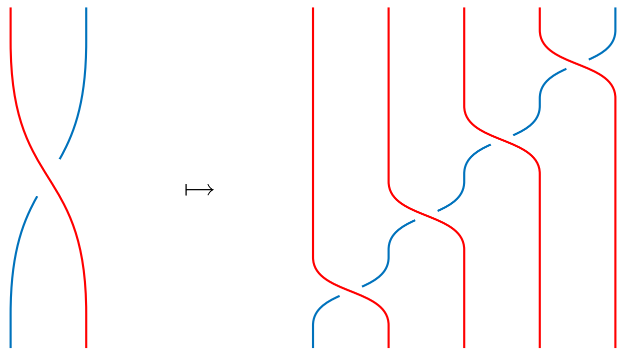
**Proposition 9.4.2.** Every cable is obtained from a map $\circ_k : B_n \times B_m \longrightarrow B_{m+n-1}$.

In general, we can define an "operadic" composition where the composition is the cabling of $n$-braids.

$$\circ_{a_1,\ldots,a_n} : B_n \times B_{a_1} \times \cdots \times B_{a_n} \longrightarrow B_{a_1+\cdots+a_n}$$

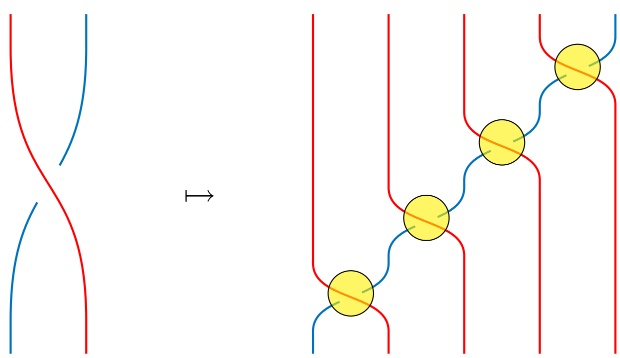
We'll want to show that this does form an operad. But before we do that we'll need to obtain an algebraic expression, based on the generators of the braids being cabled, which describe the resultant braid.

Towards that goal, consider the generator $\sigma_1$, which simply swaps the first strand over the second. Suppose we would like to substitute 4 parallel strands in the first strand of $\sigma_1$, and just one strand in the second strand of $\sigma_1$. How do we calculate this braid?



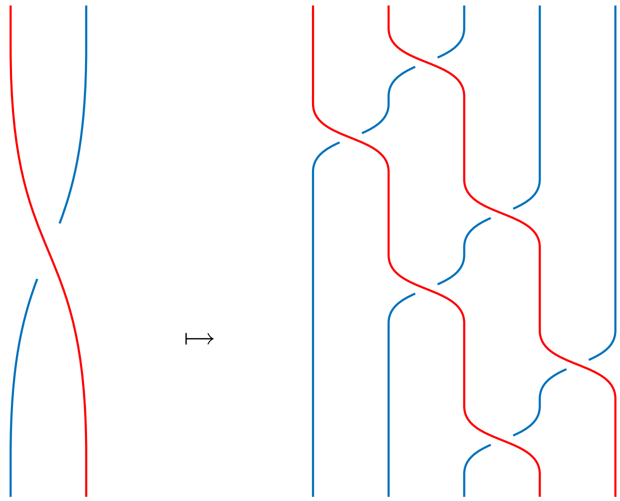
*Above is the output of $\sigma_1(4,1)$, i.e. when $k_1 = 4$ and $k_2 = 1$.*

The blue line travels diagonally down, going *underneath* each red strand once. The blue line crossing underneath the $i$-th red strand can be represented as $\sigma_i$. We then multiply all of these together to get the braid.



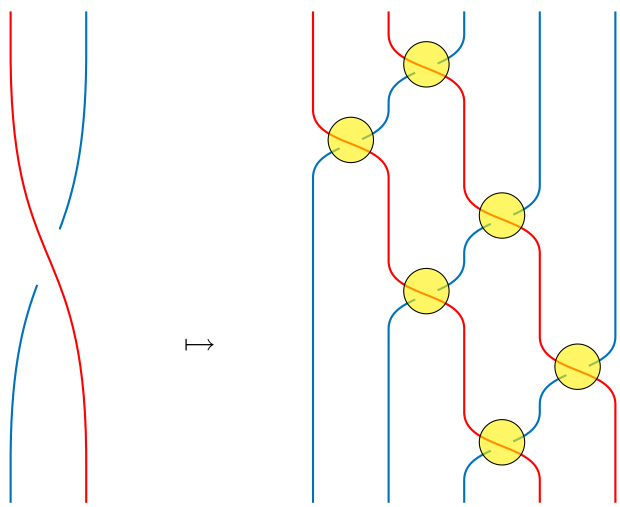
Hence we see that the braid is simply $\sigma_4\sigma_3\sigma_2\sigma_1$.

Suppose now that we would like to substitute 2 parallel strands into the first strand of $\sigma_1$, and also substitute 3 parallel strands in the second strand of $\sigma_2$. Then this produces a braid of 5 strands.



*Above is the output of $\sigma_1(2,3)$, i.e. when $k_1 = 2$ and $k_3 = 3$.*

How do we calculate this braid? Observe that the $i$-th red strand crossing over the $j$-th strand can be represented as $\sigma_{i+j-1}$. In the previous situation, $j$ was equal to 1, so it each crossing was just $\sigma_i$.

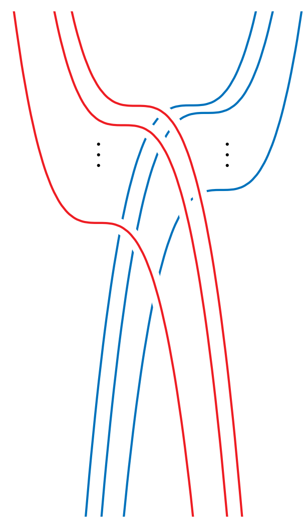


Overall, we can simply see that the braid is given by

$$(\sigma_2\sigma_1)(\sigma_3\sigma_2)(\sigma_4\sigma_3).$$

Now suppose more generally that we have $k_1$-many red lines and $k_2$-many blue lines. Then we can iteratively describe their crossings one line at time, just like we did above. The crossings will look somewhat like this:

To describe this braid, we note that there will be $k_1 \cdot k_2$-many crossings, and hence $k_1 \cdot k_2$-many generators required to describe the crossings. If we follow the first blue line, and track each time it crosses with the red lines, we see that their crossings will be $\sigma_{k_1}, \sigma_{k_1-1}, \dots, \sigma_1$. Moving onto the second blue and again traveling down, the crossings will be $\sigma_{k_1+1}, \sigma_{k_1}, \dots, \sigma_2$. If we have $k_2$-many blue lines, this will be done $k_2$ many times.



Hence we have that

$$\sigma_1(k_1, k_2) = \prod_{m=1}^{k_2} \sigma_{(k_1+m-1)}\sigma_{(k_1+m-2)} \cdots \sigma_{m} \tag{9.2}$$

where starting from $m = 1, 2, \dots, k_2$ represents us following the $m$-th blue line and recording its crossings with the red lines.

We get a similar story if we instead consider $\sigma_1^{-1}(k_1, k_2)$. Here, we are swapping $k_1$ many strands *under* $k_2$ many strands, so, we have to swap $k_1$ and $k_2$. This then gives us the expression

$$\sigma_1^{-1}(k_1, k_2) = \prod_{m=1}^{k_1} \sigma^{-1}_{(k_1-m)+1}\sigma^{-1}_{(k_1-m)+2} \cdots \sigma^{-1}_{(k_1-m)+k_2}$$

Now it is easily to generalize this to the other generators; we simply shift the indices.

$$\begin{aligned} \sigma_i(k_1, k_2) &= \prod_{m=1}^{k_2} \sigma_{(k_1+m-1+(i-1))}\sigma_{(k_1+m-2)+(i-1)} \cdots \sigma_{m+(i-1)} \\ &= \prod_{m'=i}^{(i-1)+k_2} \sigma_{(k_1+m'-1)}\sigma_{(k_1+m'-2)} \cdots \sigma_{m'} \end{aligned}$$

where we set $m' = m + (i-1)$ to reindex. Note that this returns the original formula we had once we set $i = 1$.

Thus we have that:

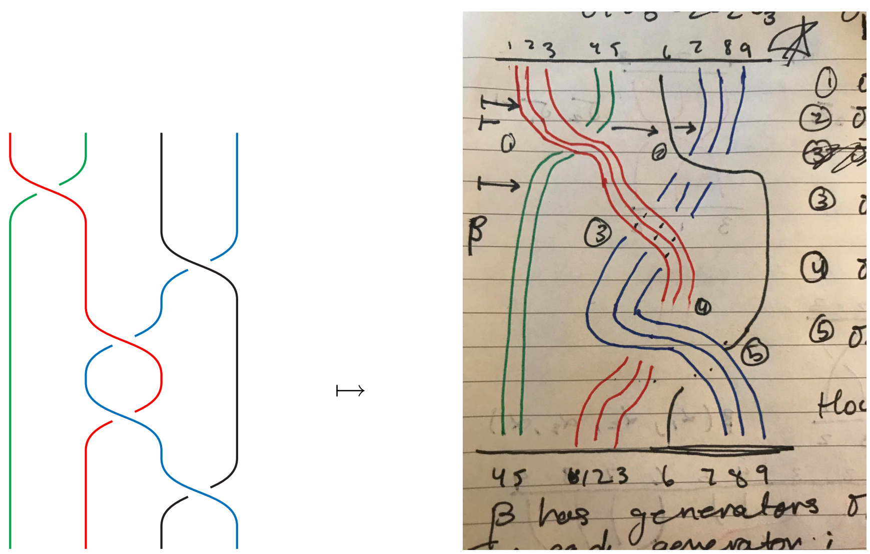
**Lemma 9.4.3.** Let $\sigma_i$ be a generator. Then the braid obtained by cabling $k_1$-many parallel lines into the $i$-th strand and $k_2$-many parallel lines into the $(i+1)$-th strand returns a braid in $B_{k_1+k_2}$ which may be expressed as

$$\sigma_i(k_1, k_2) = \prod_{m'=i}^{(i-1)+k_2} \sigma_{(k_1+m'-1)}\sigma_{(k_1+m'-2)} \cdots \sigma_{m'}$$

Now we move onto the more difficult question: suppose we have a general braiding $\beta$ of $n$ strands, and suppose we have $k_1, \dots, k_n$ sets of parallel strands. Suppose that we'd like to substitute $k_1$-parallel strands in the first strand of $\beta$, $k_2$-parallel strands in the second, all the way to $k_n$ strands in the $n$-th strand. This then defines a braid of $(k_1 + \cdots + k_n)$-many strands which we denote as

$$\beta(k_1, k_2, \dots, k_n).$$

For example, if $\beta = \sigma_1\sigma_3\sigma_2\sigma_2$, then we have $\beta$ below on the bottom left. On the bottom right, we have $\beta(k_1, k_2, k_3, k_4)$ where $k_1 = 3, k_2 = 2, k_3 = 1, k_4 = 3$.




*Above is the output of* $\beta(3, 2, 1, 3)$.

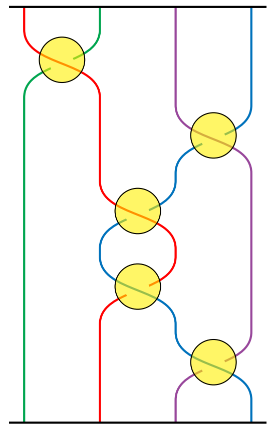
Staring at the diagram, we can see that it may be expressed as

$$(\sigma_3\sigma_2\sigma_1 \cdot \sigma_4\sigma_3\sigma_2)(\sigma_6\sigma_7\sigma_8)(\sigma_5\sigma_4\sigma_3 \cdot \sigma_6\sigma_5\sigma_4 \cdot \sigma_7\sigma_6\sigma_5)$$
$$(\sigma_5\sigma_4\sigma_3 \cdot \sigma_6\sigma_5\sigma_4 \cdot \sigma_7\sigma_6\sigma_5)(\sigma_8\sigma_7\sigma_6).$$

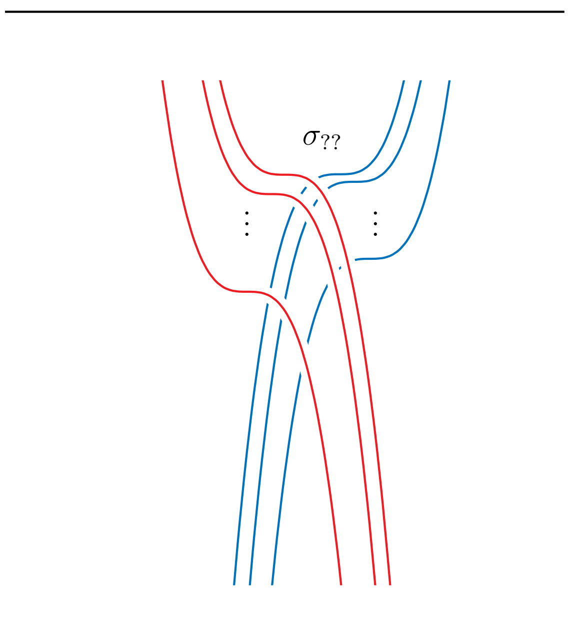
But how can we do this in general? To explain, first suppose

$$\beta = \sigma_{i_1}\sigma_{i_2} \cdots \sigma_{i_k}.$$

To draw the cabled braid $\beta(k_1, k_2, \dots, k_n)$, we see that we have $k$-crossings to focus on; these are where the crossings will happen in our cabled braid. For example, in the braid we provided above, we can highlight the crossings in yellow.



At each crossing, we're going to have something like this:




That is, at each crossing, there will be a number of red strands crossing over blue strands. If we can just describe each of these crossings using generators $\sigma_j$ like we did before, then we can describe the whole braid.

We now face the main problem. To describe an arbitrary crossing, we need to know which generators $\sigma_1, \sigma_2, \dots, \sigma_{k_1+\cdots+k_n}$ to use, and in general it's not clear which ones to use. For example, how do we describe the first crossing? We don't know, so we'll write $\sigma_{??}$. If, however, we know that the first red strand is, say the $k$-th strand in $\beta(k_1, \dots, k_n)$, then we can write the crossing as $\sigma_k$. Then we can travel down the blue line, writing $\sigma_{k-1}, \sigma_{k-2}, \dots$ until we've hit all the red strands. Then we can repeat this process for each blue line.

So to do this in general, we need to answer three questions:

- How far are all of our red strands from the left?
- How many red strands are there?
- How many blue strands are there?

If we can answer those three questions, then we can describe exactly what happens in terms of generators using formula (9.2).

We answer the first question:

**Definition 9.4.4.** Let $\beta \in B_n$ be a braid. Suppose $\beta$ can be written as a product of $k$-many generators $\beta = \sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_k}$ (where any $\sigma$ is equally possibly an inverse). Then we define the quantity

$$\varphi(\sigma_{i_1}\sigma_{i_2}\dots,\sigma_{i_j},s) = \begin{cases} \text{The order which strand } s \\ \text{is from the left after generators} \\ \sigma_{i_1}\sigma_{i_2}\dots,\sigma_{i_j} \text{ have been applied.} \end{cases}$$

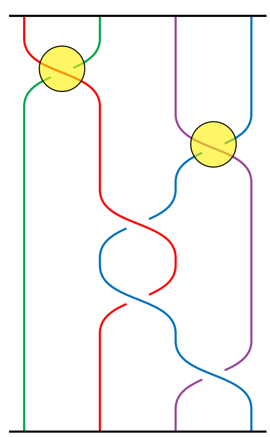
Of course, $\varphi(-,s) = s$, where $-$ represent empty input, for each strand $s$. This is because each $s$-th strand is originally the $s$-th strand.

However, a way to define this is to calculate the underlying permutation of $\sigma_i^1\sigma_j^2\dots,\sigma_k^p$ using the natural projection map $\pi : B_n \longrightarrow S_n$. Hence we see that

$$\varphi(\sigma_{i_1}\sigma_{i_2}\dots,\sigma_{i_k},s) = \pi(\sigma_{i_1}\sigma_{i_2}\dots,\sigma_{i_k})(s).$$

---

**Example 9.4.5.** Consider the braid $\sigma_1\sigma_3\sigma_2\sigma_2\sigma_3$ pictured below. Suppose we've applied $\sigma_1\sigma_3$. Then our braids are now reordered from how they were initially positioned. For instance, after the application of these generators, the green strand is now the first strand; the red strand is now the second; the blue strand is the third; and the black strand is now the fourth. Each color strand is now in a different position than which it started in.



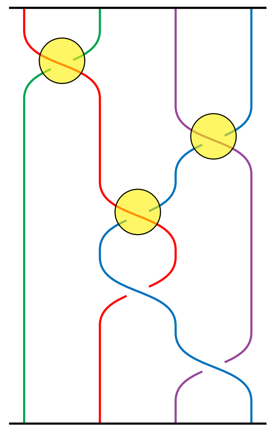
However, we can express this observation using our tool. Note that $\pi(\sigma_1\sigma_3)$ is the permutation $(1,2,3,4) \mapsto (2,1,4,3)$. Hence we see that

$$\varphi(\sigma_1\sigma_3,1) = 2 \quad \varphi(\sigma_1\sigma_3,2) = 1 \quad \varphi(\sigma_1\sigma_3,3) = 4 \quad \varphi(\sigma_1\sigma_3,4) = 3.$$

What about after the first three generators have been applied? We calculate again: $\pi(\sigma_1\sigma_3\sigma_2)$ is the permutation $(1,2,3,4) \mapsto (2,4,1,3)$. Hence we have that

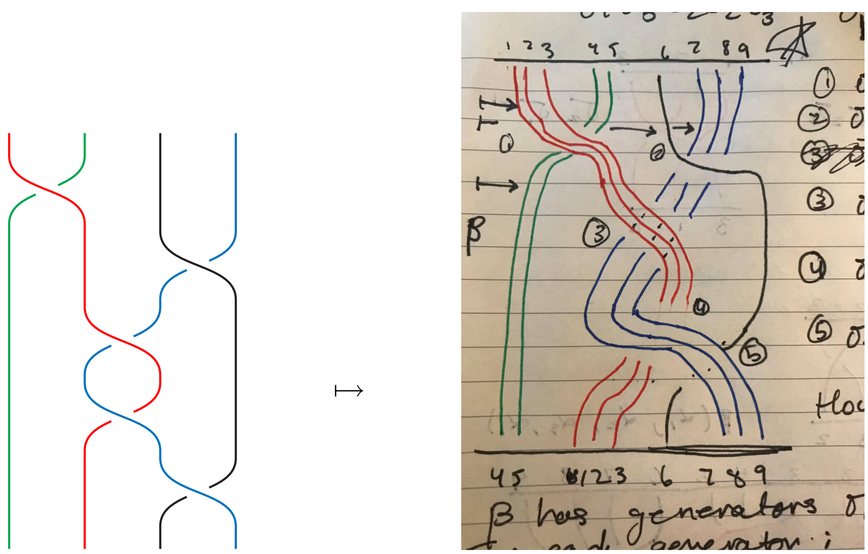
$$\varphi(\sigma_1\sigma_3\sigma_2, 1) = 2 \quad \varphi(\sigma_1\sigma_3\sigma_2, 2) = 4 \quad \varphi(\sigma_1\sigma_3\sigma_2, 3) = 1 \quad \varphi(\sigma_1\sigma_3\sigma_2, 4) = 3.$$

which matches a simple hand-count that we can perform using the picture below.



---

This tool allows us to answer our second and third questions. For example, consider again $\beta(3,2,1,3)$ where $\beta = \sigma_1\sigma_3\sigma_2\sigma_2\sigma_3$. How do we calculate, for example, the crossing ⑤, of 3 blue lines over 1 black line, as in the picture below?




*Above is the output of $\beta(3,2,1,3)$.*

This crossing is induced by $\sigma_3$, the fifth generator of $\beta$. Hence $\beta$ tells us to cross the 3nd cable over the 4rd cable. But what are these cables? From looking at the diagram, we definitely know. But in general we won't be able to just look at the diagram. However, our tool can tell us: Since we've applied $\sigma_1\sigma_3\sigma_2\sigma_2$, we see that

$$\varphi(\sigma_1\sigma_3\sigma_2\sigma_2, 3) = 4 \qquad \varphi(\sigma_1\sigma_3\sigma_2\sigma_2, 4) = 1.$$

Therefore, we're crossing blue cables over the black cables. We also now know there are $k_4 = 3$ blue cables and $k_3 = 1$ many black cables. We have almost everything we need except the following: how far are the blue cables from the left of the diagram?

Well, since the blue strands are inside of the third cable, we just need to ask how many stands are in the first and second cables. But what is the first cable? What's the second? We see that

$$\varphi(\sigma_1\sigma_3\sigma_2\sigma_2, 1) = 2. \quad \varphi(\sigma_1\sigma_3\sigma_2\sigma_2, 2) = 1.$$

Hence there are

$$k_2 + k_1 = 2 + 3 = 5$$

strands before the blue strands. We can now calculate the crossings:

$$\begin{aligned}\sigma_{5+3}\sigma_{5+2}\sigma_{5+1} &= \prod_{m=1+5}^{1+5} \sigma_{3+(m-1)}\sigma_{3+(m-2)}\sigma_m \\ &= \prod_{m=p}^{p+(r-1)} \sigma_{q+(m-1)}\sigma_{q+(m-2)}\sigma_m\end{aligned}$$

where

$$p = 1 + \overbrace{k_2 + k_3}^{\#\text{ of strands before the red strands}} \qquad q = \underbrace{k_4}_{\#\text{ of strands in the 3rd cable}} \qquad r = \overbrace{k_3}^{\#\text{ of strands in the 4th cable}}$$

Therefore we propose the following.

**Lemma 9.4.6.** Let $\beta \in B_n$ be a braid, and suppose it may be expressed as $\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_k}$ in terms of $k$-many generators. Let $k_1, \dots, k_n$ be positive integers. Then we have that

$$\beta(k_1, k_2, \dots, k_n) = \psi_1\psi_2\dots\psi_k$$

where, depending on if $\sigma_{i_j}$ is an instance of an inverse or not, we have

$$\prod_{m=p_j}^{p_j+(r_j-1)} \sigma_{q_j+(m-1)}\sigma_{q_j+(m-2)}\cdots\sigma_m \quad \text{or} \quad \prod_{m=p_j}^{p_j+(r_j-1)} \sigma^{-1}_{(q_j+m)-1}\sigma^{-1}_{(q_j-m)-2}\cdots\sigma^{-1}_m$$

where in both cases

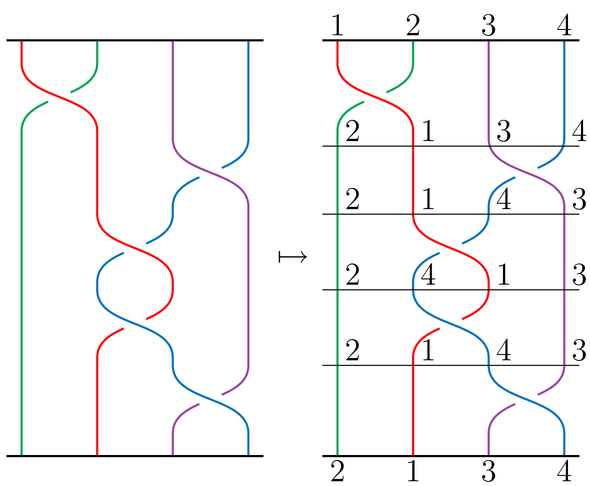
$$p_j = 1 + \overbrace{\sum_{u=1}^{i_j-1} k_{\varphi(\sigma_{i_1}\cdots\sigma_{i_{j-1}},u)}}^{\#\text{ strands before } i_j\text{-th cable}} \qquad \underbrace{q_j = k_{\varphi(\sigma_{i_1}\cdots\sigma_{i_{j-1}},i_j)}}_{\#\text{ of strands in the } i_j\text{-th cable}} \qquad \overbrace{r_j = k_{\varphi(\sigma_{i_1}\cdots\sigma_{i_{(j-1)}},(i_j+1))}}^{\#\text{ of stands in the } (i_j+1)\text{-th cable}}$$

The three quantities are the three answers to our original questions:

- After applying $\sigma_{i_1} \ldots \sigma_{i_{j-1}}$, how many strands come before the cable $i_j$, relative to the left? $p_j$.
- How many strands are in the $i_j$-th cable after applying $\sigma_{i_1} \ldots \sigma_{i_{j-1}}$? $q_j$.
- How many strands are in the $(i_j+1)$-th after applying $\sigma_{i_1} \ldots \sigma_{i_{j-1}}$? $r_j$.

---

**Example 9.4.7.** We can apply this to our previous example. Recall that $\beta = \sigma_1\sigma_3\sigma_2\sigma_2\sigma_3$. One way to interpret out braid diagram is as a sequence of permutations. In this case we see that we get five permutations because we have five generators.




First we compute the table

| $j$ | $i_j$ | $p_j$ | $q_j$ | $r_j$ |
|---|---|---|---|---|
| 1 | 1 | 1 | $k_1 = 3$ | $k_2 = 2$ |
| 2 | 3 | $1 + k_1 + k_2 = 6$ | $k_3 = 1$ | $k_4 = 3$ |
| 3 | 2 | $1 + k_2 = 3$ | $k_1 = 3$ | $k_4 = 3$ |
| 4 | 2 | $1 + k_2 = 3$ | $k_4 = 3$ | $k_1 = 3$ |
| 5 | 3 | $1 + k_1 + k_2 = 6$ | $k_4 = 3$ | $k_3 = 1$ |

This then gives us the product

$$\left(\prod_{m=p_1}^{p_1+(r_1-1)} \sigma_{q_1+(m-1)}\sigma_{q_1+(m-2)}\cdots\sigma_m\right)\left(\prod_{m=p_2}^{p_2+(r_2-1)} \sigma_{q_2+(m-1)}\sigma_{q_2+(m-2)}\cdots\sigma_m\right)$$
$$\left(\prod_{m=p_3}^{p_3+(r_3-1)} \sigma_{q_3+(m-1)}\sigma_{q_3+(m-2)}\cdots\sigma_m\right)$$
$$\left(\prod_{m=p_4}^{p_4+(r_4-1)} \sigma_{q_4+(m-1)}\sigma_{q_4+(m-2)}\cdots\sigma_m\right)\left(\prod_{m=p_5}^{p_5+(r_5-1)} \sigma_{q_5+(m-1)}\sigma_{q_5+(m-2)}\cdots\sigma_m\right)$$

which becomes

$$\left(\prod_{m=1}^{1+(2-1)} \sigma_{3+(m-1)}\sigma_{3+(m-2)}\cdots\sigma_m\right)\left(\prod_{m=6}^{6+(3-1)} \sigma_{1+(m-1)}\sigma_{1+(m-2)}\cdots\sigma_m\right)$$
$$\left(\prod_{m=3}^{3+(3-1)} \sigma_{3+(m-1)}\sigma_{3+(m-2)}\cdots\sigma_m\right)$$
$$\left(\prod_{m=3}^{3+(3-1)} \sigma_{3+(m-1)}\sigma_{3+(m-2)}\cdots\sigma_m\right)\left(\prod_{m=6}^{6+(1-1)} \sigma_{3+(m-1)}\sigma_{3+(m-2)}\cdots\sigma_m\right)$$

which reduces to

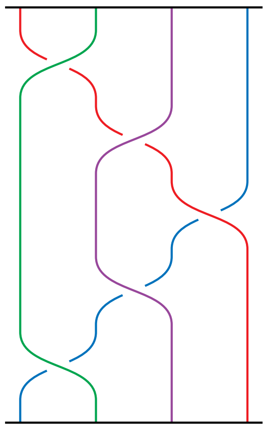
$$(\sigma_3\sigma_2\sigma_1\cdot\sigma_4\sigma_3\sigma_2)(\sigma_6\sigma_7\sigma_8)(\sigma_5\sigma_4\sigma_3\cdot\sigma_6\sigma_5\sigma_4\cdot\sigma_7\sigma_6\sigma_5)$$
$$(\sigma_5\sigma_4\sigma_3\cdot\sigma_6\sigma_5\sigma_4\cdot\sigma_7\sigma_6\sigma_5)(\sigma_8\sigma_7\sigma_6)$$

which correctly matches what we had before.

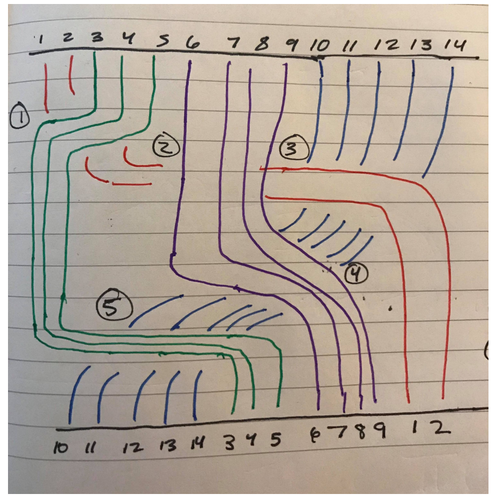
---

---

**Example 9.4.8.** We haven't looked at a braid with an under crossing. So, consider the braid $\beta = \sigma_1^{-1}\sigma_2^{-1}\sigma_3\sigma_2\sigma_1 \in B_4$, and let $k_1 = 2, k_2 = 3, k_3 = 4, k_4 = 5$. We'll want to calculate the braid $\beta(2,3,4,5)$. Below is $\beta$ and $\beta(2,3,4,5)$.



$\mapsto$




To calculate the resulting braid we need to create our table of values. This is more easily done by generating the permutation table on the left; it tells us how our cables are swapped around.

| Generator | Permutation |
|---|---|
| $\varnothing$ | $(1,2,3,4)$ |
| $\sigma_1^{-1}$ | $(2,1,3,4)$ |
| $\sigma_1^{-1}\sigma_2^{-1}$ | $(2,3,1,4)$ |
| $\sigma_1^{-1}\sigma_2^{-1}\sigma_3$ | $(2,3,4,1)$ |
| $\sigma_1^{-1}\sigma_2^{-1}\sigma_3\sigma_2$ | $(2,4,3,1)$ |
| $\sigma_1^{-1}\sigma_2^{-1}\sigma_3\sigma_2\sigma_1$ | $(4,2,3,1)$ |

| $j$ | $i_j$ | $p_j$ | $q_j$ | $r_j$ |
|---|---|---|---|---|
| 1 | 1 | 1 | $k_1 = 2$ | $k_2 = 3$ |
| 2 | 2 | $1 + k_2 = 4$ | $k_1 = 2$ | $k_3 = 4$ |
| 3 | 3 | $1 + k_2 + k_3 = 8$ | $k_1 = 2$ | $k_4 = 5$ |
| 4 | 2 | $1 + k_2 = 4$ | $k_3 = 4$ | $k_4 = 5$ |
| 5 | 1 | 1 | $k_2 = 3$ | $k_4 = 5$ |

This then generates the products

$$\left(\prod_{m=1}^{3} \sigma_{m+2}^{-1}\sigma_m^{-1}\right)\left(\prod_{m=4}^{7} \sigma_{(m+2)-1}^{-1}\sigma_m^{-1}\right)\left(\prod_{m=8}^{12} \sigma_{(m+2)-1}\sigma_m\right)\left(\prod_{m=4}^{8} \sigma_{(m+4)-1}\sigma_{(m+4)-2}\sigma_{(m+4)-3}\sigma_m\right)$$
$$\left(\prod_{m=1}^{5} \sigma_{(m+3)-1}\sigma_{(m+3)-2}\sigma_m\right)$$

which becomes

$$(\sigma_2^{-1}\sigma_1^{-1}\cdot\sigma_3^{-1}\sigma_2^{-1}\cdot\sigma_4^{-1}\sigma_3^{-1})(\sigma_5^{-1}\sigma_4^{-1}\cdot\sigma_6^{-1}\sigma_5^{-1}\cdot\sigma_7^{-1}\sigma_6^{-1}\cdot\sigma_8^{-1}\sigma_7^{-1})$$
$$(\sigma_9\sigma_8\cdot\sigma_{10}\sigma_9\cdot\sigma_{11}\sigma_{10}\cdot\sigma_{12}\sigma_{11}\cdot\sigma_{13}\sigma_{12})(\sigma_7\sigma_6\sigma_5\sigma_4\cdot\sigma_8\sigma_7\sigma_6\sigma_5\cdot\sigma_9\sigma_8\sigma_7\sigma_6\cdot\sigma_{10}\sigma_9\sigma_8\sigma_7\cdot\sigma_{11}\sigma_{10}\sigma_9\sigma_8)$$
$$(\sigma_3\sigma_2\sigma_1\cdot\sigma_4\sigma_3\sigma_2\cdot\sigma_5\sigma_4\sigma_3\cdot\sigma_6\sigma_5\sigma_4\cdot\sigma_7\sigma_6\sigma_5)$$

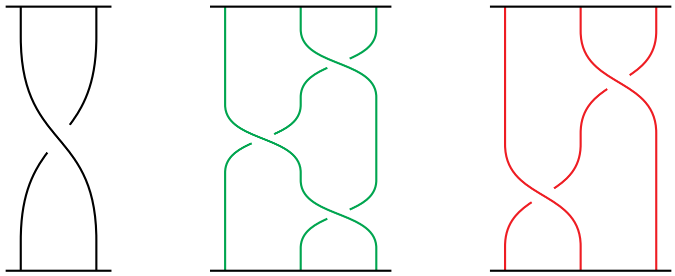
which is the correct description of the braid $\beta(2,3,4,5)$.

---

Now we can finally answer our desired question:

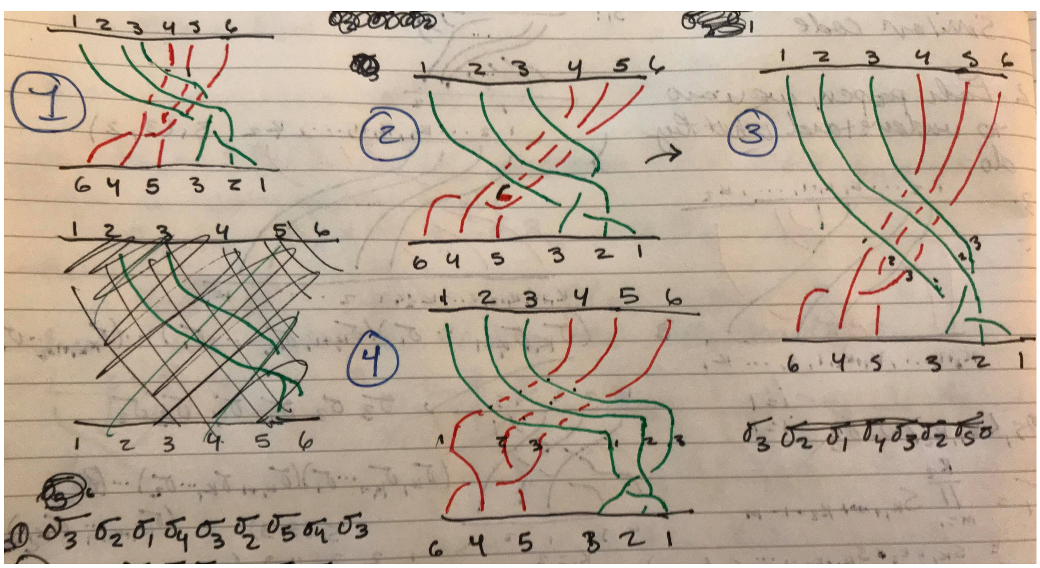
Given a braid $\beta \in B_n$, and $n$ other braids $\alpha_1 \in B_{a_1}, \ldots, \alpha_n \in B_{a_n}$, what is the formula for $\beta(\alpha_1, \ldots, \alpha_n)$?

To answer this question, we build on our previous work by making the following observation. Suppose we want to compute $\sigma_1(\alpha_1, \alpha_2)$ where $\sigma_1, \alpha_1, \alpha_2$ appear as below.

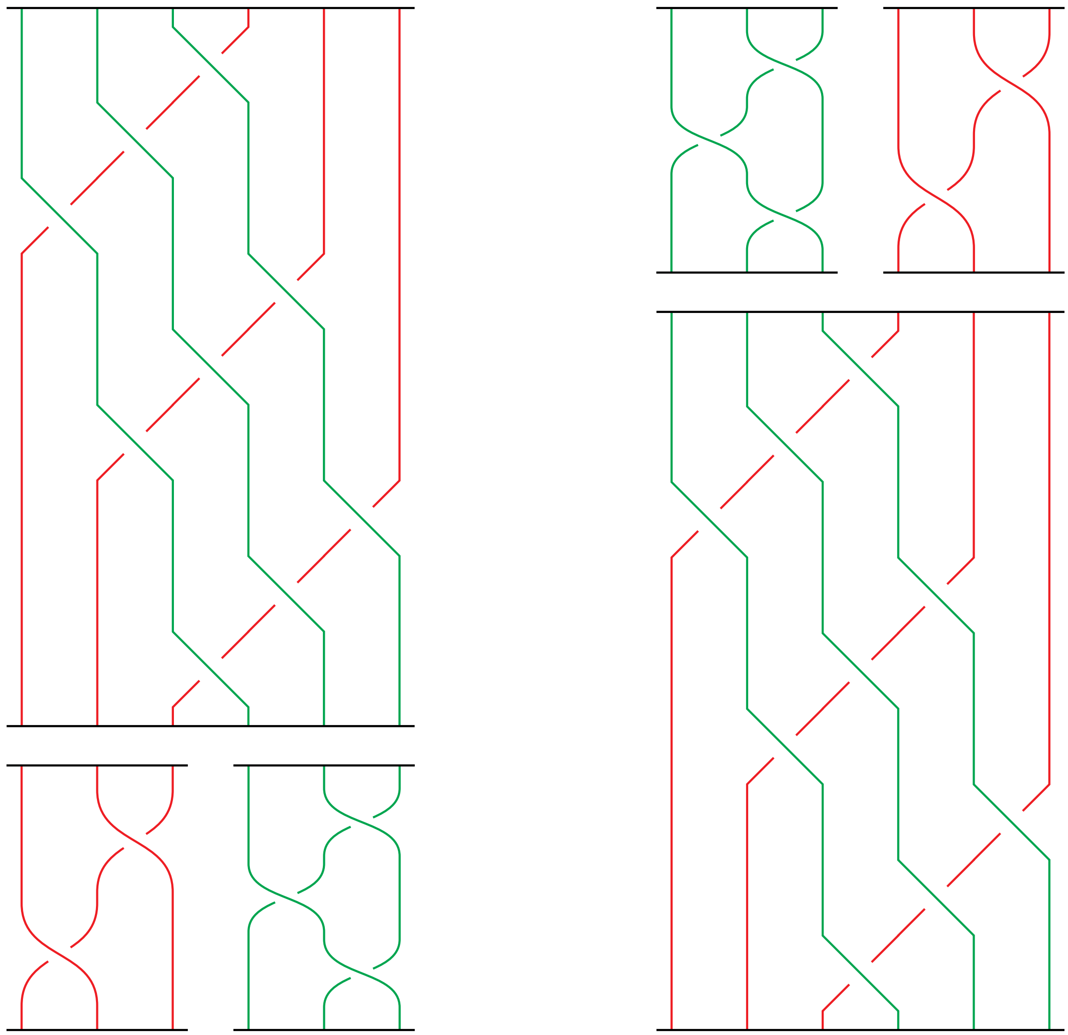


*Here we have* $\sigma_1$*,* $\alpha_1 = \sigma_2\sigma_1\sigma_2$ *and* $\alpha_2 = \sigma_2\sigma_1$*.*

Then we get the braid diagram as in ①.




However, we can all isotopies to stretch the braid to ②, then ③, and then reaching a final stage of ④. But note that ④ may be expressed in either of the equivalent ways:

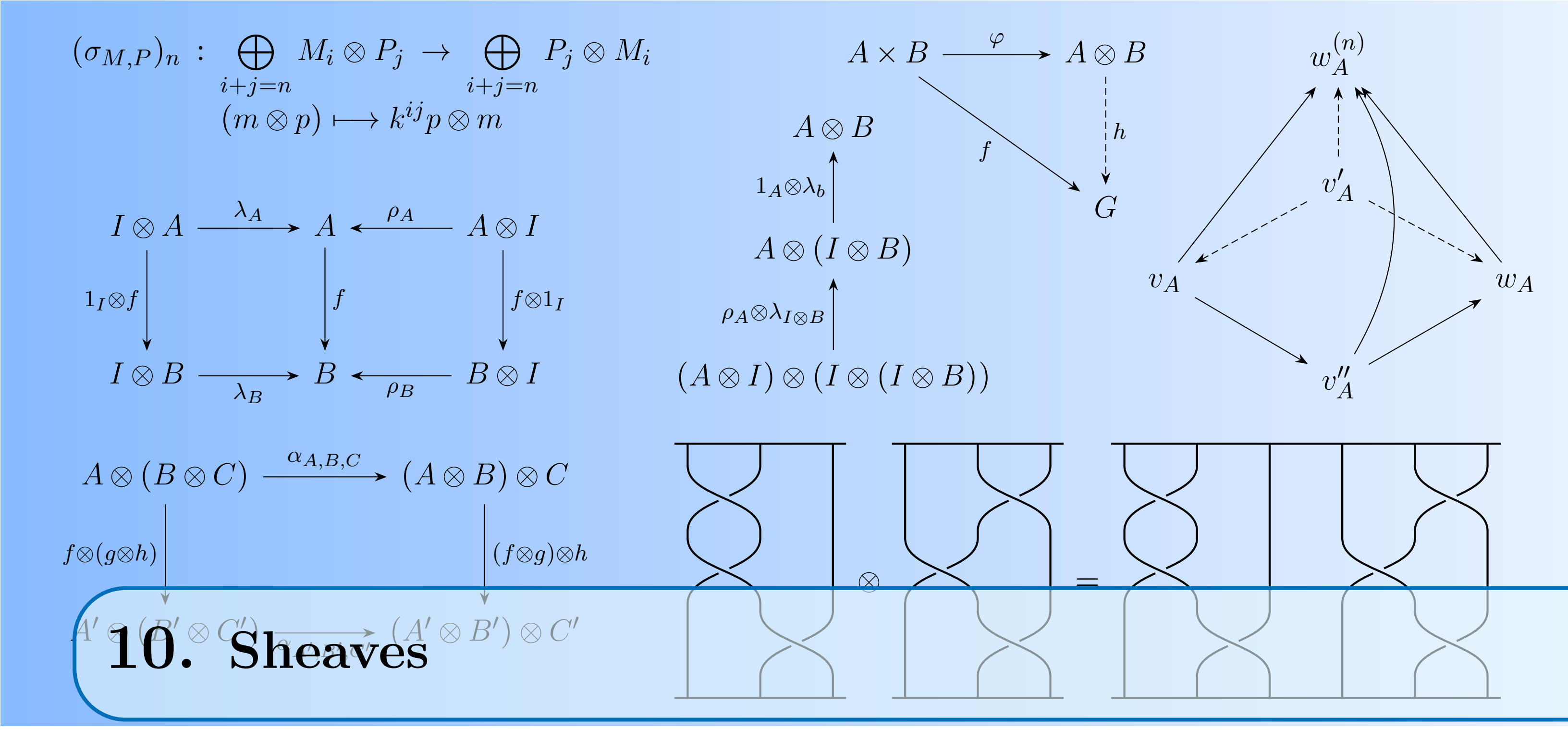


This then gives us the following idea. Suppose we want to calculate $\beta(\alpha_1, \ldots, \alpha_n)$ where $\alpha_i \in B_{a_i}$. Define $\alpha_1 \oplus \cdots \oplus \alpha_n$ as the $(a_1 + \cdots + a_n)$-braid. Suppose that $\alpha_j = \sigma_{j,i_j}, \ldots, \sigma_{j,i_{k_j}}$. Then

$$\alpha_1 \oplus \alpha_2 \oplus \cdots \oplus = (\sigma_{1,i_1}\sigma_{1,i_2}, \ldots, \sigma_{1,i_{k_1}})(\sigma_{2,(i_1+a_1)}\sigma_{2,(i_2+a_2)}, \ldots, \sigma_{2,(i_{k_1}+a_1)})$$
$$\cdots(\sigma_{n,(i_1+a_1+\cdots+a_{n-1})}\sigma_{2,(i_2+a_2)}, \ldots, \sigma_{n,(i_{k_1}+a_1+\cdots+a_{n-1})})$$

which concatenates the braid horizontally. Then we see that

$$\beta(\alpha_1, \ldots, \alpha_n) = \beta(a_1, a_2, \ldots, a_n) \circ \alpha_1 \oplus \alpha_2 \oplus \cdots \oplus \alpha_n.$$



# 10. Sheaves

## 10.1 Topological Presheaves and Sheaves

Let $X$ be a topological space. Denote the set of open subsets of $X$ as $\mathbf{Open}(X)$. We can impose the structure of a thin category on this set by declaring that, for two open sets $U$ and $V$,

$$\mathrm{Hom}_{\mathbf{Open}(X)}(U, V) = \begin{cases} \{\bullet\} & \text{if } U \subseteq V \\ \varnothing & \text{otherwise} \end{cases}$$

That is, we allow a single morphism from $U$ to $V$ if and only if $U \subseteq V$. Now suppose $Y$ is another topological space. Then for each open subset $U$ of $X$ we may construct the set

$$C(U) = \{f : U \longrightarrow Y \mid f \text{ is continuous }\}.$$

Observe that if $U \subseteq V \subseteq X$ are open sets, then there is function

$$\rho^V_U : C(V) \longrightarrow C(U)$$

where each $f : V \longrightarrow Y$ is mapped to its restriction $f|_U : U \longrightarrow Y$. What follows is an important observation: If we have a chain of three open subsets $U \subseteq V \subseteq W$, then any continuous function $f : W \longrightarrow Y$ can be restricted to $f|_V : V \longrightarrow Y$, which can then be restricted to $f|_V|_U : U \longrightarrow Y$. However, we obtain the same result if we instead just restrict $f$ to $U$ in the first place. That is, $f|_V|_U = f|_U$. In our notation, this implies that

$$\rho^W_V \circ \rho^V_U = \rho^W_U.$$

What we have on our hands is a *contravariant* functor (since the relation $U \subseteq V$ induces a function $C(V) \longrightarrow C(U)$). As covariant functors are easier to think about, we can equivalently express this as a covariant functor:

$$C : \mathbf{Open}(X)^{\text{op}} \longrightarrow \mathbf{Set}$$

which is an example of the concept of a *presheaf*.

**Definition 10.1.1.** A **presheaf (of sets on a topological space $X$)** is a covariant functor $F : \mathbf{Open}(X)^{\text{op}} \longrightarrow \mathbf{Set}$. We spell out the details: A presheaf consists of

**(PS1)** an assignment of open sets $U \subseteq X$ to sets $F(U)$

**(PS2)** a function $\rho^V_U : F(V) \longrightarrow F(U)$ whenever $U \subseteq V$ such that

  **(Identity)** $\rho^U_U : F(U) \longrightarrow F(U)$ is the identity

  **(Composition)** $\rho^W_V \circ \rho^V_U = \rho^W_U$ whenever $U \subseteq V \subseteq W$

A **morphism of presheaves** is a natural transformation between presheaves.

A few comments are to be made about this definition.

- *About* **Set**. The codomain of a presheaf doesn't have to be **Set**. Usually, the value of our presheaves are sets of functions, but sometimes such sets have additional structure. Therefore, the codomain could be **Ab**, **Ring**, or another category where the objects are sets plus some mathematical structure. In these cases, we'd obtain a **presheaf of abelian groups**, a **presheaf of rings**, and so forth.
- *About the naming.* The only reason this is called a presheaf is because, as the reader may guess, this idea is a precursor to the concept of a sheaf.
- The fact that we can formulate morphisms between presheaves prompts us to define the **category of presheaves (of sets) on $\mathcal{C}$** which we denote as $\mathbf{Psh}(X, \mathbf{Set})$.

We now offer some examples of presheaves. The examples we offer will be topological presheaves, i.e., presheaves on $\mathbf{Open}(X)$ for some topological space $X$. This is because many interesting and useful examples of presheaves appear in this way. This is also done so that we can offer our first definition of sheaf with as littel confusion as possible.

---

**Example 10.1.2.** Consider the introductory example of this section, and instead take $Y = \mathbb{R}$. Then in this case,

$$C(U) = \{f : U \longrightarrow \mathbb{R} \mid f \text{ is continuous}\}.$$

However, observe that $C(U)$ is actually an $\mathbb{R}$-module: if $f, g : U \longrightarrow \mathbb{R}$ are continuous, then so is $f + g : U \longrightarrow \mathbb{R}$. Moreover, if $a \in \mathbb{R}$, then $a \cdot f : U \longrightarrow \mathbb{R}$ is continuous. These operations satisfy the criteria for $C(U)$ to be an $\mathbb{R}$-module. Therefore, when $Y = \mathbb{R}$, we obtain a presheaf on $\mathbb{R}$-**Mod**, and we may write

$$C : \mathbf{Open}(X)^{\text{op}} \longrightarrow \mathbb{R}\text{-}\mathbf{Mod}.$$

We will return to this example later on.

**Example 10.1.3.** For every open set $D$ of the complex plane $\mathbb{C}$, define the set

$$H(D) = \{f : D \longrightarrow \mathbb{C} \mid f \text{ is holomorphic. }\}$$

Observe that this induces a functor $H : \mathbf{Open}(\mathbb{C})^{\mathrm{op}} \longrightarrow \mathbf{Set}$, and hence we have a presheaf of sets. Moreover, this is actually a $\mathbb{C}$-module, so that what we have is actually a presheaf of $\mathbb{C}$-modules; hence we write $H : \mathbf{Open}(\mathbb{C})^{\mathrm{op}} \longrightarrow \mathbb{C}\text{-}\mathbf{Mod}$.

**Example 10.1.4.** Let $X$ be a topological space, and consider the functor $B : \mathbf{Open}(X)^{\mathrm{op}} \longrightarrow \mathbb{R}\text{-}\mathbf{Mod}$, defined as follows. For an open subset $U \subseteq X$, we define

$$B(U) = \{f : U \longrightarrow \mathbb{R} \mid f \text{ is bounded}\}.$$

By bounded, we mean that $f : U \longrightarrow \mathbb{R}$ is bounded if there exists a constant $M \in \mathbb{R}$ such that, for all $x \in U$, $|f(x)| \leq M$. This example becomes important later, specifically in that it is an example of a presheaf which is not a sheaf (yet to be defined).

Our next goal is to offer our first definition of a sheaf. To motivate the definition, we will consider our introductory example.

Recall our presheaf $C : \mathbf{Open}(X)^{\mathrm{op}} \longrightarrow \mathbf{Set}$. Consider an open set $U$ with an open cover $\mathcal{U} = \{U_i\}_{i\in\lambda}$. Then every $f : U \longrightarrow \mathbb{R}$ in $C(U)$ corresponds to an element of $F(U_i)$ for all $i$; it is simply the restriction $f|_{U_i} \longrightarrow \mathbb{R}$.

A natural question is the converse: If I have such an open cover $\mathcal{U}$ of $U$, and a family of continuous functions $f_i : U_i \longrightarrow Y$, is there a continuous function $f : U \longrightarrow \mathbb{R}$ such that $f|_{U_i} = f_i$ for all $i$?

Immediately, the answer is no: simply take a family in which the functions disagree on their overlaps. Thus, what if our family does agree on their overlaps? This would mean that, for every pair $i, j$,

$$f_i|_{U_i\cap U_j} = f_j|_{U_i\cap U_j}.$$

(Of course, $U_i \cap U_j$ could be empty; but we don't know in general, so we just play it safe and consider *all* pairs $i, j \in \lambda$.) The answer now is affirmative, there is an fact a $f : U \longrightarrow \mathbb{R}$ where $f|_{U_i} = f_i$ for all $i$. Thus we see that $C : \mathbf{Open}(X)^{\mathrm{op}} \longrightarrow \mathbf{Set}$ is a rather special type of presheaf, and we call this kind of functor a sheaf.

**Definition 10.1.5.** Let $X$ be a topological space. A **topological sheaf (of sets) on** $X$ is a presheaf $F : \mathbf{Open}(X)^{\mathrm{op}} \longrightarrow \mathbf{Set}$ such that, for every open set $U$ and any open cover $\mathcal{U} = \{U_i\}_{i\in\lambda}$

of $U$, the following two properties hold:

**(SH1)** If $f,\, g \in F(U)$ are such that $f|_{U_i} = g|_{U_i}$ for all $i \in \lambda$, then $f = g$.

**(SH2)** Suppose that for all $i \in \lambda$, we have $h_i \in F(U_i)$ such that $h_i|_{U_i \cap U_j} = h_j|_{U_i \cap U_i}$ (i.e., a family of $h_i$ which agree on all possible overlaps). Then there exists a $h \in F(U)$ such that $h|_{U_i} = h_i$ for all $i \in \lambda$.

A few comments about this definition:

- In our definition, **SH2** is our main axiom of focus. We add **SH1** so that the given $h \in F(U)$ in **SH2** is necessarily unique.
- Once again, the codomain of our sheaf does not have to **Set**. We will see this in a few examples.
- With the notion of a morphism of sheaves, we can define the **category of topological sheaves (of Sets)**, denoted $\mathbf{Sh}(X, \mathbf{Set})$, to be the category with objects sheaves and morphisms with natural transformations.

We end this definition by defining a **morphism of sheaves**; it is simply a natural transformation between sheaves.

We now offer a few examples of topological sheaves.

---

**Example 10.1.6.** Consider again the introductory example $C : \mathbf{Open}(X)^{\text{op}} \longrightarrow \mathbb{R}\text{-}\mathbf{Mod}$. We show that this is a sheaf. Towards that goal, let $U$ be an open with open cover $\mathcal{U} = \{U_i\}_{i \in \lambda}$.

**(SH1)** Suppose $f, g : U \longrightarrow \mathbb{R}$ are continuous functions which agree on the overlaps of the open cover. Then in this case it's clear that $f = g$.

**(SH2)** Suppose $f_i : U_i \longrightarrow \mathbb{R}$ is a family of continuous functions such that $f_i|_{U_i \cap U_j} = f_j|_{U_i \cap U_j}$ for all $i, j \in \lambda$. Construct a function $\varphi : U \longrightarrow \mathbb{R}$ pointwise as follows: Given a $p \in U$, there exists some $k \in \lambda$ such that $p \in U_k$. Therefore, let $\varphi(p) = f_k(p)$; agreement on overlaps makes this well defined.
We show that $\varphi$ is continuous. For an open set $V$ of $\mathbb{R}$, define $\varphi^{-1}(V) = \bigcup_{i \in \lambda} f_i^{-1}(V)$. As this is a union of open sets, $\varphi^{-1}(V)$ is open and hence $\varphi$ is continuous.

As we've satisfied **SH1** and **SH2**, we see that this is a sheaf.

---

A reader familiar with topology will note that our work towards the axiom **SH2** in the last example is nothing more than the standard proof of the *Pasting Lemma* from topology.

---

**Example 10.1.7.** Consider the presheaf $H : \mathbf{Open}(\mathbb{C}) \longrightarrow \mathbb{C}\text{-}\mathbf{Mod}$ which sends open sets of $\mathbb{C}$ to the $\mathbb{C}$-module of holomorphic functions defined on them.

This is also a sheaf, which we verify. Let $U$ be an open set of $\mathbb{C}$ and $\mathcal{U} = \{U_i\}_{i \in \lambda}$ an open cover.

**(SH1)** This is true in the same was as the last example.

**(SH2)** Let $f_i : U_i \longrightarrow \mathbb{C}$ be a family of holomorphic functions such that each $f_i$ agree on all possible overlaps. Define $f : U \longrightarrow \mathbb{C}$ in the obvious way. To show that $f$ is holomorphic on $U$, pick any $p \in U$. Then $p \in U_k$ for some $k$, and hence there exists an open set $D_k$ of $p$ such that $f_k(z) = \sum_{n=1}^{\infty} a_n(z-p)^n$, i.e., it has a power series representation. This then gives us a well-defined power series representation for $f$, so that $f$ is holomorphic.

---

We now offer an example of a presheaf which is not a sheaf.

---

**Example 10.1.8.** Consider the presheaf $B : \mathbf{Open}(X) \longrightarrow \mathbb{R}\text{-}\mathbf{Mod}$ where $B(U)$ is the set of all bounded functions $f : U \longrightarrow \mathbb{R}$.

In general, this is not a sheaf; axiom **SH2** is usually broken. For example, take $X = \mathbb{R}$, and consider the open set $(0,1)$, with the open cover given by the sets $\{U_i = \left(\frac{1}{i}, 1\right) \mid i = 1, 2, \dots\}$. Observe that the functions

$$f_i(x) : \left(\frac{1}{i}, 1\right) \longrightarrow \mathbb{R} \qquad f_i(x) = \frac{1}{x}$$

agree on their overlaps, but clearly there is no bounded function $f : (0,1) \longrightarrow \mathbb{R}$ such that $f|_{U_i} = f_i$ for all $i$. Hence, this is not a sheaf.

---

## 10.2 Abstracting Sheaves

We will now take a more categorical approach to extract the key properties of a sheaf, so that we may generalize our logic. Towards that goal, we'll introduce a second definition of a sheaf, one which is equivalent to what the reader has already seen; it will offer a new perspective. To motivate this perspective, we will again use our canonical sheaf of continuous functions:

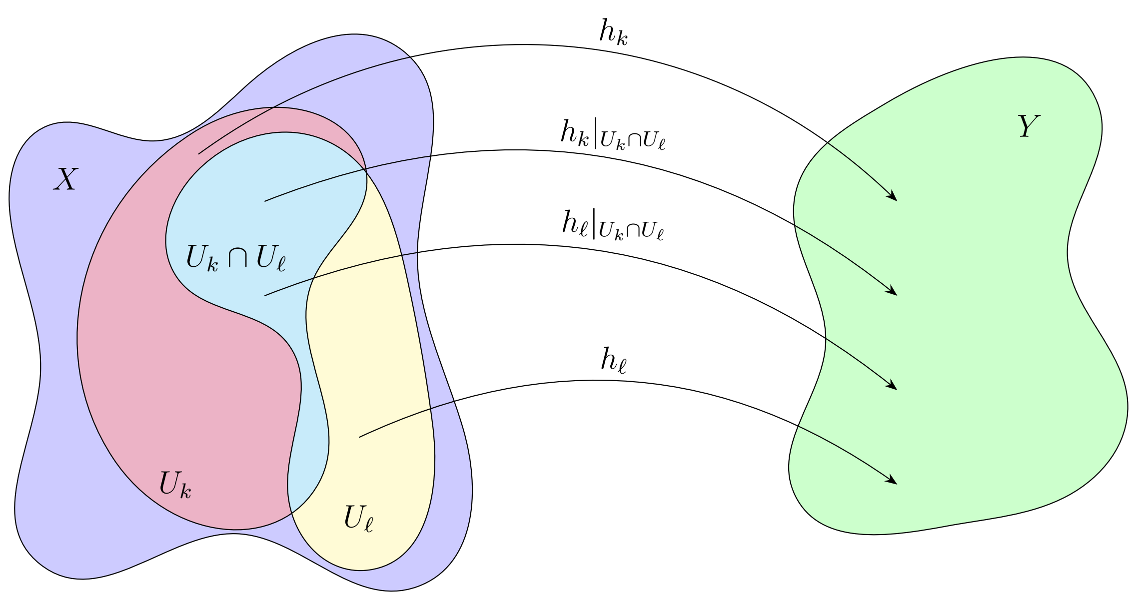
$$C : \textbf{Open}(X)^{\text{op}} \longrightarrow \textbf{Set} \qquad C(U) = \{f : U \longrightarrow Y \mid f \text{ is continuous}\}$$

Consider an open set $U$ of $X$, and let $\mathcal{U} = \{U_i\}_{i\in\lambda}$ be an open cover of $U$. Let us make a few nontrivial observations. The reader is strongly encouraged to move forward with pen and paper in hand and to draw lots of pictures.

- A family of continuous functions $h_i : U_i \longrightarrow Y$ can be viewed as an element $(h_i)_{i\in\lambda}$ of the product $\prod_{i\in\lambda} C(U_i)$.
- Using our open cover $\mathcal{U}$, we can define for each pair $k, \ell \in \lambda$ the functions

  $$p_{k,\ell}, q_{k,\ell} : \prod_{i\in\lambda} C(U_i) \longrightarrow C(U_k \cap U_\ell)$$

  where
  $$p_{k,\ell}\Big((h_i)_{i\in\lambda}\Big) = h_k\Big|_{U_k\cap U_\ell} \qquad \text{and} \qquad q_{k,\ell}\Big((h_i)_{i\in\lambda}\Big) = h_\ell\Big|_{U_k\cap U_\ell}.$$
  With a lot of notation, a picture may help.




- The fact that the functions $p_{k,\ell}, q_{k,\ell}$ exist for all $k, \ell \in \lambda$ implies the existence of $p$ and $q$ below which make the diagram commute. (This is just applying the universal property of the product $\prod_{i,j} F(U_i \cap U_j)$.) These two functions are rather important.

$$
\begin{tikzcd}
 & \displaystyle\prod_{i\in\lambda} F(U_i) \arrow[dl, "q_{k,\ell}"'] \arrow[d, dashed, shift right, "q"'] \arrow[d, dashed, shift left, "p"] \arrow[dr, "p_{k,\ell}"] & \\
F(U_k \cap U_\ell) & \displaystyle\prod_{i,j\in\lambda} F(U_i \cap U_j) \arrow[l, "\pi_{k,\ell}"] \arrow[r, "\pi_{k,\ell}"'] & F(U_k \cap U_\ell)
\end{tikzcd}
$$

Now consider the set of all $(h_i)_{i\in\lambda} \in \prod_{i\in\lambda} F(U_i)$ such that they agree on overlaps; i.e., such that $h_i\Big|_{U_i\cap U_j} = h_j\Big|_{U_i\cap U_j}$ for all $i, j \in \lambda$. We call this set $\mathrm{Eq}(p, q)$:

$$
\mathrm{Eq}(p,q) = \left\{(h_i)_{i\in\lambda} \in \prod_{i\in\lambda} F(U_i) \mid p\Big((h_i)_{i\in\lambda}\Big) = q\Big((h_i)_{i\in\lambda}\Big)\right\}.
$$

However, since $C$ is a sheaf, we know that for every such $(h_i)_{i\in\lambda}$ in $\mathrm{Eq}(p, q)$ there exists a unique $h : U \longrightarrow Y$ such that $h|_{U_i} = h_i$. Therefore, we see that

$$
\mathrm{Eq}(p,q) \cong C(U)
$$

Okay, so that's just a slightly more complicated way of expressing $C(U)$. What's interesting about this, however, is that $\mathrm{Eq}(p, q)$ is quite literally the equalizer of $p$ and $q$ (hence the naming we chose for the set).

$$
F(U) \xrightarrow{\ e\ } \prod_{i\in\lambda} F(U_i) \underset{p}{\overset{q}{\rightrightarrows}} \prod_{i,j\in\lambda} F(U_i \cap U_j).
$$

This is the motivation behind the following definition of a sheaf, which is exactly equivalent to our previous one.

**Definition 10.2.1.** A **sheaf (of sets) on a topological space** $X$ is a functor

$$
F : \mathbf{Open}(X)^{\mathrm{op}} \longrightarrow \mathbf{Set}
$$

with the following property: If $U$ is an open set and $\mathcal{U} = \{U_i\}_{i\in\lambda}$ an open cover of $U$, then $F(U)$ is an equalizer of $p$, $q$, constructed using $\mathcal{U}$ as above. The equalizer diagram is below:

$$
F(U) \xrightarrow{\ e\ } \prod_{i\in\lambda} F(U_i) \underset{p}{\overset{q}{\rightrightarrows}} \prod_{i,j\in\lambda} F(U_i \cap U_j).
$$

We remark two comments on this definition.

- It is more important to understand the *philosophy* of the above definition rather than the literal text of it (of course, that's necessary). For example, a topological space does in fact speak of families of sets which are closed under arbitrary union and finite intersection. But that's a *literal* definition, and not the philosophy of a topological space.

- There are many ways to state the definition of a sheaf. The one offered above is very powerful because it allows us to quickly capture many useful situations and it is useful for proofs.

Now before we move on, we are going to briefly introduce a new concept.

**Definition 10.2.2.** Let $\mathcal{C}$ be a category and $C$ an object of $\mathcal{C}$. A **sieve on** $C$ is a set $S$ which is a subset of all morphisms with codomain $C$:

$$S \subseteq \{f \mid f : B \longrightarrow C \text{ and } f \text{ is a morphism of } \mathcal{C}\}$$

with following property.

**(SV1)** If $f$ is in $S$, then $f \circ h$ is in $S$ for any composable $h$.

We will demonstrate an example of this concept, specifically to capture why we care about it.

---

**Example 10.2.3.** Let $X$ be a topological space, and consider the category $\mathbf{Open}(X)$. Let $U$ be an open set of $X$. To speak of a sieve on $U$, we must first realize that the set of all objects with codomain $U$ is simply the set

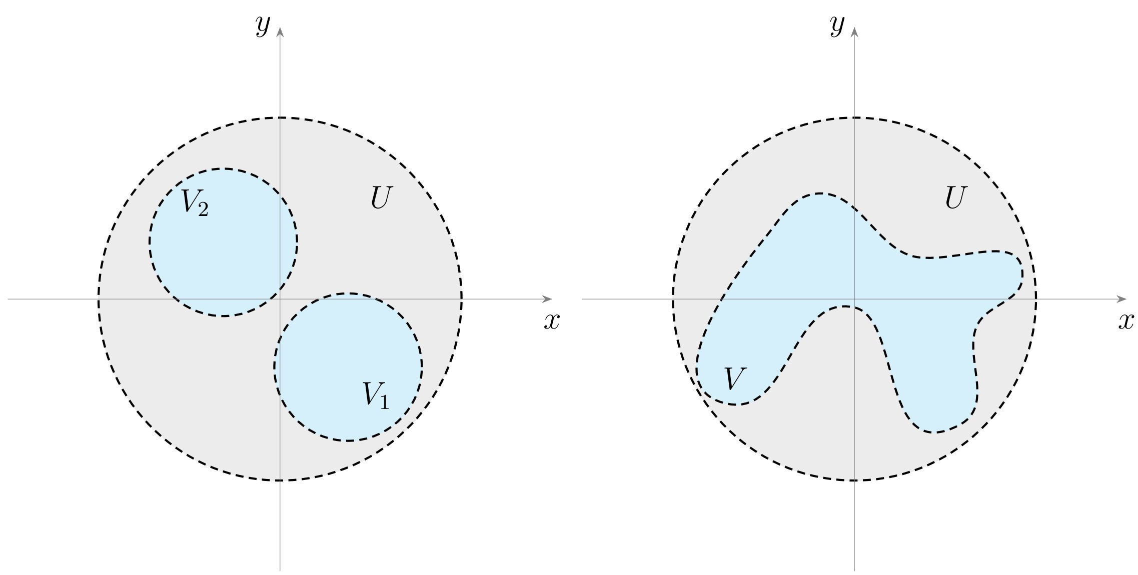
$$\Omega_U = \{V \subseteq U \mid V \text{ is open}\}$$

This set may actually be treated as the object set of the full subcategory $\mathbf{Open}(U)$ of $\mathbf{Open}(X)$.

So, what is a sieve in this case? It is any $S \subseteq \Omega_U$ such that

**(SV1)** If $V \in S$, $V'$ is open, and $V' \subseteq V$, them $V' \in S$.

Take note that this is a bit of subtle concept; it's a very versatile definition. For example, considering $\mathbb{R}^2$ with its standard topology, the following (blue) open sets create sieves on the same open set (the open disk at the origin).




On the left, we consider the set of all open sets contained in $V_1$ and $V_2$; this is a sieve on the open disk (which we call $U$ to be consistent with our notation and discussion). On the right, we consider the set of all open sets contained in the weirdly shaped $V$; this is also a sieve on $U$.

---

Some important facts about sieves on topological spaces that will be of interest to us.

- Every open set $V \subseteq U$ corresponds to a sieve, which we call a **principal sieve**. This sieve is simply the set of all open $V'$ contained in $V$. In the previous example, the weirdly shaped region inside the open disk at the origin is a principal sieve.
- Every open cover of $\mathcal{U} = \{U_i\}_{i\in\lambda}$ creates a **covering sieve** $S_{\mathcal{U}}$. This sieve is the set of all open $V$ such that $V \subseteq U_i$ for some $i$, and where $V' \subseteq V$ implies $V'$ is also in the set.

  Additionally, a covering sieve induces a (fairly stupid) functor $\mathcal{S}$, where:

$$\mathcal{S} : \textbf{Open}(X)^{\text{op}} \longrightarrow \textbf{Set} \qquad \mathcal{S}(V) = \begin{cases} \{\bullet\} & \text{If } V \in S_{\mathcal{U}} \\ \varnothing & \text{otherwise.} \end{cases}$$

We are now prepared to continue our discussion. Our goal now will be to express the equalizer $E$ of $p, q$ categorically (i.e., without reference to its elements). Let $P : \mathcal{O}(X)^{\text{op}} \longrightarrow \textbf{Set}$ be a presheaf. Given an open set $U$ with open cover $\mathcal{U} = \{U_i\}_{i\in\lambda}$, we may construct $p, q$ using $\mathcal{U}$ as before, and take their equalizer $E$:

$$E \xrightarrow{\;e\;} \prod_{i\in\lambda} P(U_i) \underset{p}{\overset{q}{\rightrightarrows}} \prod_{i,j\in\lambda} P(U_i \cap U_j).$$

We now prove the following result.

**Lemma 10.2.4.** Let $E$ be the equalizer of $p, q$ constructed using an open cover $\mathcal{U}$ of $U$. Let $\mathcal{S}$ be the sieve functor induced by $\mathcal{U}$. Then

$$E \cong \text{Hom}(\mathcal{S}, P) \quad \text{or, in alternate notation,} \quad E \cong \text{Nat}(\mathcal{S}, P)$$

That is, there is a bijection between $E$ and all natural transformations between $\mathcal{S}$ and $P$.

*Proof.* We know that

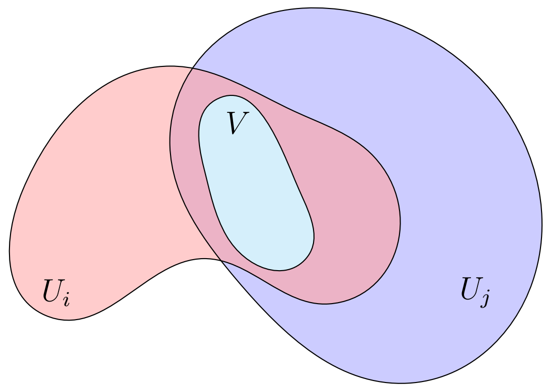
$$E = \left\{ (h_i)_{i\in\lambda} \;\middle|\; h_i|_{U_i\cap U_j} = h_j|_{U_i\cap U_j} \text{ for all } i, j \right\}.$$

We'll show that every $(h_i)_{i\in\lambda}$ can be used to build a natural transformation between $\mathcal{S} \longrightarrow P$. Showing the other direction is not hard.

Let $S_{\mathcal{U}}$ be our covering sieve induced by $\mathcal{U}$. Consider an element $(h_i)_{i\in\lambda}$ in $E$. For each $V \in S_{\mathcal{U}}$, we define $h_V \in P(V)$ as

$$h_V = h_i|_V$$

where $i$ is the index such that $V \subseteq U_i$. Of course at least one index exists, but it might not be the only index. Thus, a natural objection to this definition is the following question: What if $V$ is contained in $U_i$ and $U_j$ for distinct $i$, $j$? In this case, how do we define $h_V$?




If $(h_i)_{i \in \lambda} \in E$, then we know that agreement on the overlaps is guaranteed and so we may unambiguously write $h_i|_V = h_j|_V = h_V$. Hence, each $V \in S_{\mathcal{U}}$ corresponds to some *unique* $h_V \in P(V)$ for every $(h_i)_{i \in \lambda} \in E$. Furthermore, we know that if $V' \subseteq V$, then $h_V|_{V'} = h_{V'}$.

These facts allow us to create the following natural transformation $\theta : \mathcal{S} \longrightarrow P$ using an element $(h_i)_{i \in \lambda}$ of $E$, as follows.

- If $V \in S_{\mathcal{U}}$, we write $\theta_V : \{\bullet\} \longrightarrow P(V)$ where $\theta_V(\bullet) = h_V$, the unique $h_V$ we already know exists.

This allows us to create the function

$$\varphi : E \longrightarrow \mathrm{Hom}(\mathcal{S}, P) \qquad \varphi\Big((h_i)_{i \in \lambda}\Big) \mapsto (\theta : \mathcal{S} \longrightarrow P)$$

It is not difficult to show that every natural transformation between $\mathcal{S}$ and $P$ corresponds to a unique element in $E$, thereby giving us an inverse to this function. Thus we have our result. ■

The above result is key to the the following proposition, which is what allows us to speak of a sheaf more abstractly. Before we introduce the proposition, we make a few comments.

- Let

**Proposition 10.2.5.** Let $P : \mathbf{Open}(X)^{\mathrm{op}} \longrightarrow \mathbf{Set}$ be a presheaf. Then $P$ is a presheaf if and only if for every open set $U$ and covering sieve $S$ of $U$,

## 10.3 Stalks and Germs

Let $(I, \le)$ be a partially ordered set, and suppose we have a functor $F : I \longrightarrow \textbf{Set}$. With this functor, denote $F(i) = A_i$ and when $i \le j$, $F(i \le j) = f_{ij} : A_i \longrightarrow A_j$. The limit of this functor $\varinjlim_{i \in I} F$ will be a set $A$ equipped with functions $\varphi_i : A_i \longrightarrow A$ with the universal property displayed below.

$$\begin{tikzcd}
A_i \arrow[rr, "f_{ij}"] \arrow[dr, "\varphi_i"'] \arrow[ddr, bend right, "\psi_i"'] & & A_j \arrow[dl, "\varphi_j"] \arrow[ddl, bend left, "\psi_j"] \\
& A \arrow[d, dashed] & \\
& K &
\end{tikzcd}$$

We may naively suppose that $A = \coprod_{i \in I} A_i = \{(a, i) \mid a \in A_i, i \in I\}$, since such a set admits a family of functions $\text{inc}_i : A_i \longrightarrow \coprod_{i \in I} A_i$. However, we cannot guarantee that this the triangle

$$\begin{tikzcd}
A_i \arrow[rr, "f_{ij}"] \arrow[dr, "\text{inc}_i"'] & & A_j \arrow[dl, "\text{inc}_j"] \\
& \coprod_{i \in I} A_i &
\end{tikzcd}$$

will commute. In fact, it will never commute, since it would imply that for each $a \in A_i$, $(a, i) = (f_{ij}(a), j)$, which cannot happen as the tuples are mismatched. Since it is too strong to demand equality, we can define an equivalence relation $\sim$ on $\coprod_{i \in I} A_i$ as follows: For $i \le j$, we say $(a, i) \sim (b, j)$ if $b = f_{ij}(a)$. We can then set

$$A = \coprod_{i \in I} A_i \Big/ \sim$$

and define a family of maps $\varphi_i : A_i \longrightarrow A$ which maps each $a \in A_i$ to its equivalence class under this relation. This then allows the desired triangle to commute and satisfies the universal property necessary for it to be the limit.

We now apply this construction to our story with sheaves.

**Definition 10.3.1.** Let $X$ be a topological space and $F : \mathcal{O}(X) \longrightarrow \textbf{Set}$ a sheaf. For any point

$x \in X$, we define the **stalk of $F$ in $x$**, denoted $F_x$, as the colimit

$$\varinjlim_{x \in U} F(U).$$

The above notation is a bit informal, but many people use it, so we will stick with it and explicitly describe this limit as follows. Each point $x \in X$ induces a functor $F^{(x)} : \mathcal{O}(X)_x \longrightarrow \textbf{Set}$ where $\mathcal{O}_x$ is the category of open sets of $X$ *containing* $x$, and $F^{(x)}(U) = F(U)$. We can then more formally say

$$\varinjlim_{x \in U} F(U) = \varinjlim_{U \in \mathcal{O}(X)_x} F^{(x)}.$$

Therefore, we can say that

$$\varinjlim_{x \in U} F(U) = \coprod_{U | x \in U} F(U) \Big/ \sim$$

where $\sim$ is the equivalence relation described previously. In this instance, the equivalence relation translates as follows. Let $U_1 \subseteq U_2$ be two open sets. Then we say $(f, U_1) \sim (g, U_2)$ if $g\Big|_{U_1} = f$.

We can make this more refined as follows. Let $U_1, U_2$ be more generally any two open sets such that $V = U_1 \cap U_2 \neq \varnothing$. Then clearly $V \subseteq U_1$ and $V \subseteq U_2$. Now suppose, $(f, V) \sim (g_1, U_1)$ for some $f, g_1$, and $(f, V) \sim (g_2, U_2)$. Then we now have that

$$(g_1, U_1) \sim (g_2, U_2) \iff g_1\Big|_V = g_2\Big|_V.$$

Thus we have translated our original equivalence relation into a more useful one. To summarize, we have that our stalk is the set

$$\varinjlim_{x \in U} F(U) = \Big\{ [(f, U)] \mid x \in U \text{ open}, f \in F(U), \Big\}$$

where $(f, U)$ is a representative of its equivalence class $[(f, U)]$, described explicitly as

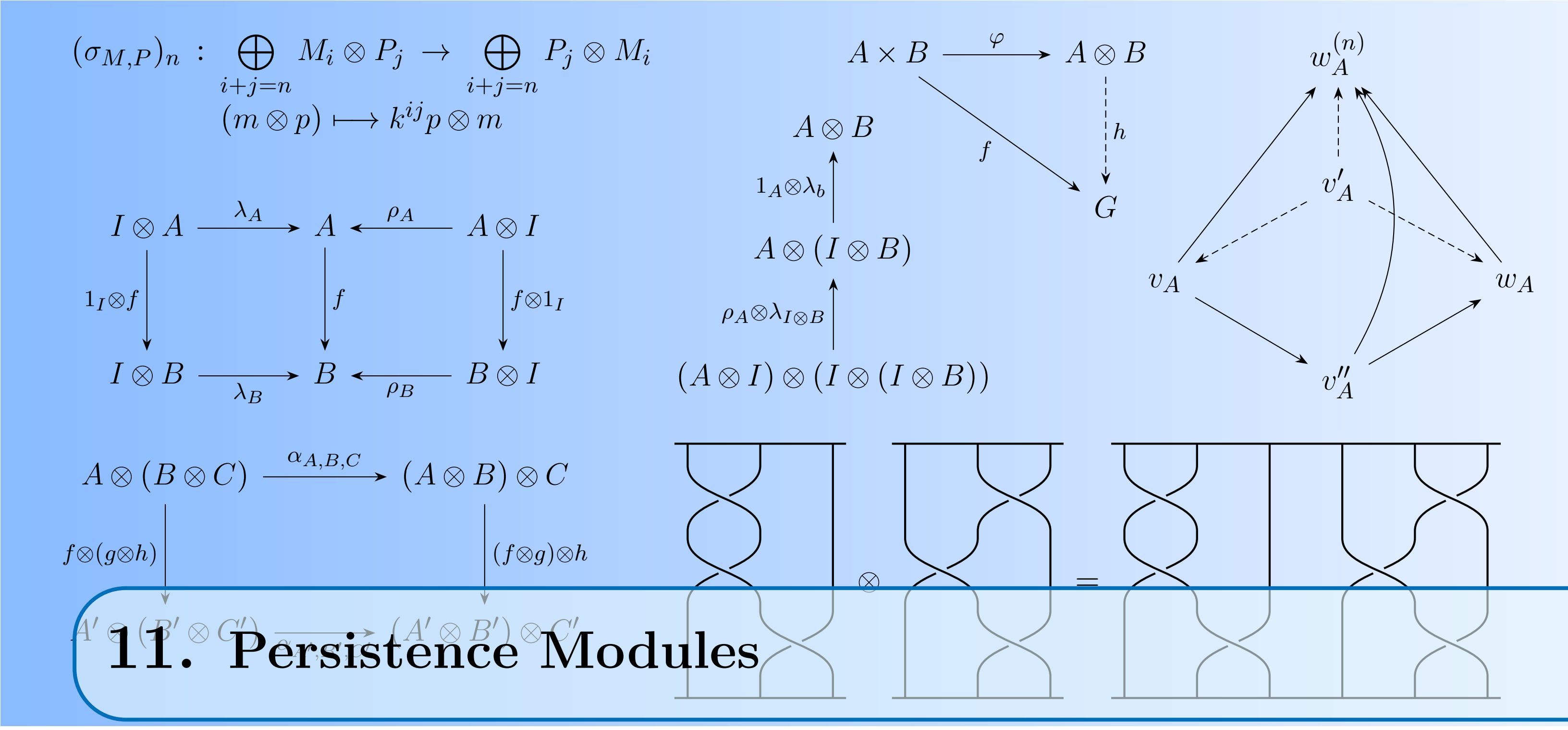
$$[(f, U)] = \Big\{ (g, V) \mid g \in F(U), x \in V \text{ open and } g\Big|_V = f\Big|_V \Big\}.$$

The above line leads to our next definition.

**Definition 10.3.2.** Let $U$ be an open set containing $x$. There naturally exists projection map

$$\pi_U : F(U) \longrightarrow F_x \qquad f \mapsto [(f, U)].$$

Therefore, for each $f \in F(U)$, we define the **germ of $f$ in $x$** to be the equivalence class $[(f, U)]$ in the stalk $F_x$.




# 11. Persistence Modules

## 11.1 Persistence modules on $\mathbb{R}$.

**Definition 11.1.1.** Let $\mathcal{C}$ be a category, and denote $(\mathbb{R}, \leq)$ to be the poset category on $\mathbb{R}$ with respect to the natural relation $\leq$. We define a functor $F : (\mathbb{R}, \leq) \longrightarrow \mathcal{C}$ to be a **persistence module**.

Thus we can say that a persistence module is an element of the functor category $\mathcal{C}^{\mathbb{R}}$.

A persistence module allows us to model the evolution of objects within some category $\mathcal{C}$. For example, if we have some ascending chain of vector spaces

$$\cdots \longrightarrow V_{i-1} \longrightarrow V_i \longrightarrow V_{i+1} \longrightarrow \cdots$$

then we say that such a chain is a persistence module since it can be modeled as a functor from $\mathbb{R} \longrightarrow \textbf{Vec}$.

Let $S = \{s_1, s_2, \dots, s_n\}$ be a finite subset of $\mathbb{R}^n$. Then we can describe an adjunction

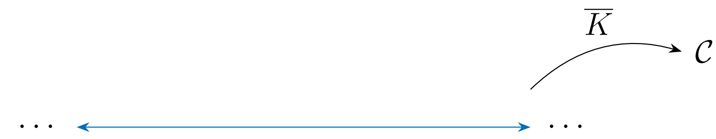
$$\mathcal{C}^{\mathbb{R}} \rightleftarrows \mathcal{C}^S$$

as follows. First observe that since $S \subseteq \mathbb{R}$, there exists a restriction functor $R : \mathcal{C}^{\mathbb{R}} \longrightarrow \mathcal{C}^S$, which acts as a restriction (hence the naming $R$):

$$R(F : \mathbb{R} \longrightarrow \mathcal{C}) = F\Big|_S : S \longrightarrow \mathcal{C}.$$

How can we write a functor going in the opposite direction? That is, given a persistence module which acts on $S$,

is there a canonical way to extend this to a persistence module which acts on the rest of $\mathbb{R}$?




One way we may extend a persistence module $K : S \longrightarrow \mathcal{C}$ in $\mathcal{C}^S$ to a persistence module in $\mathcal{C}^{\mathbb{R}}$ is to define a functor $\overline{K} : \mathbb{R} \longrightarrow \mathcal{C}$ where

$$\overline{K}(r) = \begin{cases} I & \text{if } s < s_1 \\ K(r) & \text{if } s_i \le r \le s_{i+1} \\ K(r_n) & \text{if } r > s_n \end{cases} = \begin{cases} I & \text{if } r < \min(S) \\ K(s_r) & \text{where } s_r \text{ is the largest } s_r \in \text{ S such that } s_r \le r. \end{cases}$$

Now consider a morphism $\eta : K \longrightarrow P$ in $\mathcal{C}^S$; that is, a natural transformation. By our above procedure we have a way of discussing the objects $\overline{K}$ and $\overline{P}$; but can we obtain a natural transformation $\overline{\eta} : \overline{K} \longrightarrow \overline{P}$ from $\eta$? That is, may we extend this relationship to a functor?

First, observe that we may write $\eta : K \longrightarrow P$ as follows.

$$\begin{tikzcd} P(s_1) \arrow[r] & P(s_2) \arrow[r] & \cdots \arrow[r] & P(s_{n-1}) \arrow[r] & P(s_n) \\ K(s_1) \arrow[r] \arrow[u, "\eta_{S_1}"] & K(s_2) \arrow[r] \arrow[u, "\eta_{S_2}"] & \cdots \arrow[r] & K(s_{n-1}) \arrow[r] \arrow[u, "\eta_{S_{n-1}}"] & K(s_n) \arrow[u, "\eta_{S_n}"'] \end{tikzcd}$$

The top and bottom rows come about by functoriality of $K$ and $P$, while the upward arrows are the family of morphisms created by the existence of a natural transformation.

We can extend this to a natural transformation $\overline{\eta} : \overline{K} \longrightarrow \overline{P}$ by stating

$$\overline{\eta}_r = \begin{cases} 1_I & \text{if } r < s_1, \text{ where } I \text{ is initial} \\ \eta_{s_r} & \text{where } s_r \text{ is the largest } s_r \in S \text{ such that } s_r \le r. \end{cases}$$

## Adjoint Functors

Thus we see that we really do have a functor $\mathcal{C}^S \longrightarrow \mathcal{C}^{\mathbb{R}}$ on our hands If we denote this as a functor $E : \mathcal{C}^S \longrightarrow \mathcal{C}^{\mathbb{R}}$, where $E$ can be read as *extends*, then we overall have

$$\mathcal{C}^{\mathbb{R}} \underset{E}{\overset{R}{\rightleftarrows}} \mathcal{C}^S \ .$$

We can now demonstrate that this pair of functors gives rise to an adjunction; there a few ways to do this. We'll demonstrate that

$$\mathrm{Hom}_{\mathcal{C}^S}(K, P_S) \cong \mathrm{Hom}_{\mathcal{C}^{\mathbb{R}}}(\overline{K}, P)$$

is natural, where $P_S = \mathrm{R}(P)$ and $\overline{K} = E(K)$. Towards this goal, consider a morphism $\eta : K \longrightarrow P_S$. Then we have something like this again

$$\begin{tikzcd}
P_s(s_1) \arrow[r] & P_s(s_2) \arrow[r] & \cdots \arrow[r] & P_s(s_{n-1}) \arrow[r] & P_s(s_n) \\
K(s_1) \arrow[r] \arrow[u, "\eta_1"] & K(s_2) \arrow[r] \arrow[u, "\eta_2"] & \cdots \arrow[r] & K(s_{n-1}) \arrow[r] \arrow[u, "\eta_{n-1}"] & K(s_n) \arrow[u, "\eta_n"']
\end{tikzcd}$$

Now we seek a natural transformation $\eta' : \overline{K} \longrightarrow P$. Since $\overline{K}$ is constructed from $K$, a good choice would be to write $\eta'_{s_i} = \eta_{s_i}$ for $s_i \in S$. Now our concern is considering how to define $\eta'_r$ when $r \notin S$. That is, we want something like

$$\begin{tikzcd}
\cdots \arrow[r] & P(s_i) \arrow[r] & P(r) \arrow[r] & P(s_{i+1}) \arrow[r] & \cdots \\
\cdots \arrow[r] & \overline{K}(s_i) \arrow[r] \arrow[u, "\eta'_{s_i}"] & \overline{K}(r) \arrow[r] \arrow[u, "\eta'_r"] & \overline{K}(s_{i+1}) \arrow[r] \arrow[u, "\eta'_{s_{i+1}}"] & \cdots
\end{tikzcd}$$

To define the morphism in red, we first recall that in this situation we have $K(r) = K(s_i)$. Hence we know that any morphism from $K(r)$ must originate from $K(s_i)$; one such morphism we already know about is $\eta_{s_i} : K(s_i) \longrightarrow P_s(s_i)$. Now, $P_s(s_i) = P(s_i)$; and in our case the desired target for $\eta'$ is $P(r)$, not $P(s_i)$. However, we can compose this with the morphism $P(j) : P(s_i) \longrightarrow P(r)$. where $j : s_i \longrightarrow r$.

$$\begin{tikzcd}
\cdots \arrow[r] & P(s_i) \arrow[r, "j"] & P(r) \arrow[r] & P(s_{i+1}) \arrow[r] & \cdots \\
\cdots \arrow[r] & \overline{K}(s_i) \arrow[r, equal] \arrow[u, "\eta'_{s_i}"] & \overline{K}(r) \arrow[r] \arrow[u, "\eta'_r"] & \overline{K}(s_{i+1}) \arrow[r] \arrow[u, "\eta'_{s_{i+1}}"] & \cdots
\end{tikzcd}$$

Therefore, in this case we define

$$\eta'_r := P(j) \circ \eta_{s_i}.$$

which necessarily forces commutativity, and hence demonstrating naturality of $\eta'$. Now what if $r < s_1$ or $s_n < s$? In the first case, $K(r) = I$, and $\eta'_r$ becomes the unique morphism from $I \longrightarrow P(r)$. This presents one benefit of adding the criteria $K(r) = I$ if $r < s_1$. By uniqueness of this morphism we get a commutative square. In the second case, we proceed as above. Therefore

$$\eta'_r = \begin{cases} i_{P(r)} : I \longrightarrow P(r) & \text{if } r < s_1 \\ P(j : s_i \longrightarrow r) \circ \eta_{s_i} & \text{where } s_i \text{ is the largest } s \in S \text{ such that } s \leq r. \end{cases}$$

Therefore, we can define a map $\varphi : \mathrm{Hom}_{\mathcal{C}^S}(K, P_S) \longrightarrow \mathrm{Hom}_{\mathcal{C}^{\mathbb{R}}}(\overline{K}, P)$ where

$$\varphi(\eta : K \longrightarrow P_S) = \eta' : \overline{K} \longrightarrow P.$$

Consider the map $\psi : \text{Hom}_{\mathcal{C}^{\mathbb{R}}}(\overline{K}, P) \longrightarrow \text{Hom}_{\mathcal{C}^S}(K, P_S)$ where

$$\psi(\sigma : \overline{K} \longrightarrow P) = \sigma' : K \longrightarrow P_S$$

where we set $\sigma'_s = \sigma_s$. While this map is particularly boring, we're discussing it because we can now see that $\psi$ and $\varphi$ are inverses of each other. Therefore, we see that we have a bijection between the hom-sets, as desired.

## Naturality.

Finally, we must demonstrate naturality. So suppose we have a natural transformation $\alpha : K \longrightarrow K'$ between two persistence modules $K, K' : S \longrightarrow \mathcal{C}$. Consider the squares below, which we do not yet know commutes.

$$\begin{tikzcd}
\text{Hom}_{\mathcal{C}^S}(K, P_S) \arrow[r, "\varphi"] \arrow[d] & \text{Hom}_{\mathcal{C}^{\mathbb{R}}}(\overline{K}, P) \arrow[d] \\
\text{Hom}_{\mathcal{C}^S}(K', P_S) \arrow[r, "\varphi"] & \text{Hom}_{\mathcal{C}^{\mathbb{R}}}(\overline{K'}, P)
\end{tikzcd}
\qquad
\begin{tikzcd}
\eta : K \longrightarrow P_S \arrow[r, mapsto] \arrow[d, mapsto] & \eta' : \overline{K} \longrightarrow P \arrow[d, mapsto] \\
\eta \circ \alpha : K' \longrightarrow P_S \arrow[r, mapsto] & \begin{array}{c} \eta' \circ \overline{\alpha} : K' \longrightarrow P \\ = \\ (\eta \circ \alpha)' : K' \longrightarrow P \end{array}
\end{tikzcd}$$

Note that on one hand,

$$\overline{\alpha}_r = \begin{cases} 1_I & \text{if } r < s_1, \text{ where } I \text{ is initial} \\ \alpha_{s_r} & \text{where } s_r \text{ is the largest } s_r \in S \text{ such that } s_r \leq r. \end{cases}$$

and

$$\eta'_r = \begin{cases} i_{P(r)} : I \longrightarrow P(r) & \text{if } r < s_1 \\ P(j : s_i \longrightarrow r) \circ \eta_{s_i} & \text{where } s_i \text{ is the largest } s \in S \text{ such that } s \leq r. \end{cases}$$

so that

$$\begin{aligned}
(\eta' \circ \overline{\alpha})_r &= \begin{cases} i_{P(r)} : I \longrightarrow P(r) & \text{if } r < s_1 \\ \Big(P(j : s_i \longrightarrow r) \circ \eta\Big) \circ \alpha & \text{where } s_r \text{ is the largest } s_r \in S \text{ such that } s_r \leq r. \end{cases} \\
&= \begin{cases} i_{P(r)} : I \longrightarrow P(r) & \text{if } r < s_1 \\ P(j : s_i \longrightarrow r) \circ (\eta \circ \alpha) & \text{where } s_r \text{ is the largest } s_r \in S \text{ such that } s_r \leq r. \end{cases} \\
&= (\eta \circ \alpha)'_r.
\end{aligned}$$

Since we know that $\Big(P(j : s_i \longrightarrow r) \circ \eta\Big) \circ \alpha = P(j : s_i \longrightarrow r) \circ (\eta \circ \alpha)$. Thus we see that the previous squares we discussed do in fact commute.

Now suppose we have a natural transformation $\sigma : P \longrightarrow P'$ between two functors $P, P' : \mathbb{R} \longrightarrow \mathcal{C}$. Consider the diagrams below, which we will show are commutative.

$$\begin{tikzcd}
\mathrm{Hom}_{\mathcal{C}^S}(K, P_S) \arrow[r, "\varphi"] \arrow[d] & \mathrm{Hom}_{\mathcal{C}^{\mathbb{R}}}(\overline{K}, P) \arrow[d] \\
\mathrm{Hom}_{\mathcal{C}^S}(K, P'_S) \arrow[r, "\varphi"] & \mathrm{Hom}_{\mathcal{C}^{\mathbb{R}}}(\overline{K}, P')
\end{tikzcd}
\qquad
\begin{tikzcd}
\eta : K \longrightarrow P_S \arrow[r, mapsto] \arrow[d, mapsto] & \eta' : \overline{K} \longrightarrow P \arrow[d, mapsto] \\
\sigma' \circ \eta : K \longrightarrow P'_S \arrow[r, mapsto] & \begin{matrix} \sigma \circ \eta' : K \longrightarrow P' \\ = \\ (\sigma' \circ \eta)' : K \longrightarrow P' \end{matrix}
\end{tikzcd}$$

To show this, observe that

$$\begin{aligned}
\sigma \circ \eta' &= \begin{cases} \sigma_r \circ i_{P(r)} : I \longrightarrow P'(r) & \text{if } r < s_1 \\ \sigma_r \circ P(j : s_i \longrightarrow r) \circ \eta_{s_i} & \text{where } s_i \text{ is the largest } s \in S \text{ such that } s \leq r. \end{cases} \\
&= \begin{cases} i_{P'(r)} : I \longrightarrow P'(r) & \text{if } r < s_1 \\ P'(j : s_i \longrightarrow r) \circ (\sigma \circ \eta)_{s_i} & \text{where } s_i \text{ is the largest } s \in S \text{ such that } s \leq r. \end{cases} \\
&= \begin{cases} i_{P'(r)} : I \longrightarrow P'(r) & \text{if } r < s_1 \\ P'(j : s_i \longrightarrow r) \circ (\sigma' \circ \eta)_{s_i} & \text{where } s_i \text{ is the largest } s \in S \text{ such that } s \leq r. \end{cases} \\
&= (\sigma' \circ \eta)'.
\end{aligned}$$

The diagrams below can assist to seeing why this is the case. First, the change in purple occurs by commutativity of the diagram on the left; the commutativity results due to the universal nature of morphisms originating from the initial object $I$. Second, the changes in green and red occur by commutativity of the diagram on the right.

$$\begin{tikzcd}
P(r) \arrow[rr, "\sigma_r"] & & P'(r) \\
& I \arrow[ul, "i_{P(r)}"] \arrow[ur, "i_{P'(r)}"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
P'(s_i) \arrow[r, "P'(j)"] & P'(r) \\
P(s_i) \arrow[u, "\sigma_{s_i} = \sigma'_{s_i}"] \arrow[r, "P(j)"] & P'(r) \arrow[u, "\sigma_r"'] \\
K(s_i) \arrow[u, "\eta_{s_i}"] \arrow[r, equal] & \overline{K}(r) \arrow[u, "\eta'_r"']
\end{tikzcd}$$

Thus we see that our original squares are commutative. At this point, we can conclude that we do in fact have an adjunction

$$\begin{tikzcd}
\mathcal{C}^{\mathbb{R}} \arrow[r, shift left, "R"] & \mathcal{C}^S \arrow[l, shift left, "E"]
\end{tikzcd}$$

as desired.

## 11.2 Generalized Persistence Modules.

**Definition 11.2.1.** Let $P$ be a preorder. Then a **generalized persistence module** is a functor $F : P \longrightarrow \mathcal{D}$.

Therefore, we may view $D^P$ to be the category of generalized persistence modules on $P$.

**Definition 11.2.2.** A **translation** on $P$ is a functor $\Gamma : P \longrightarrow P$ such that $x \leq \Gamma(x)$ for all $x$. Equivalently, it is any functor such that there exists a natural transformation $\eta_\Gamma : I \longrightarrow \Gamma$.

We can denote the category of translations on $P$ as $\mathbf{Trans}_P$. Note that this is a preorder. Since $P$ is a preorder, any two natural transformations between two functors must necessarily be equal. Moreover, every pair of translations must have a natural transformation; that is, one (or both) of the diagrams below must commute for any $x \leq y$ in $P$.

$$\begin{tikzcd}
\Gamma(x) \arrow[r] \arrow[d] & \Gamma(y) \arrow[d] \\
K(x) \arrow[r] & K(y)
\end{tikzcd}
\qquad
\begin{tikzcd}
K(x) \arrow[r] \arrow[d] & K(y) \arrow[d] \\
\Gamma(x) \arrow[r] & \Gamma(y)
\end{tikzcd}.$$

Thus we set $\Gamma \leq K$ whenever there exists a natural transformation $\eta_{\Gamma K} : \Gamma \longrightarrow K$.

**Definition 11.2.3.** Let $P$ be a preorder and $\Gamma, K \in \mathbf{Trans}_P$. Suppose $F, G \in \mathcal{D}^P$. We say $F, G$ are $(\Gamma, K)$-interleaved if there exists a pair of natural transformations $\varphi : F \longrightarrow G \circ \Gamma$ and $\psi : G \longrightarrow F \circ K$ such that

$$\begin{tikzcd}
F(x) \arrow[r] \arrow[d, "\varphi_x"'] & F(y) \arrow[d, "\varphi_y"] \\
G(\Gamma(x)) \arrow[r] & G(\Gamma(y))
\end{tikzcd}
\qquad
\begin{tikzcd}
G(x) \arrow[r] \arrow[d, "\psi_x"'] & G(y) \arrow[d, "\psi_y"] \\
F(K(x)) \arrow[r] & F(K(y))
\end{tikzcd}$$

$$\begin{tikzcd}
F(x) \arrow[rr, "F(\eta_{K(\Gamma(x))})"] \arrow[dr, "\varphi_x"'] & & F(K(\Gamma(x))) \\
& G(\Gamma(x)) \arrow[ur, "\psi_{\Gamma(x)}"'] &
\end{tikzcd}
\quad
\begin{tikzcd}
G(x) \arrow[rr, "G(\eta_{\Gamma(K(x))})"] \arrow[dr, "\psi_x"'] & & G(\Gamma(K(x))) \\
& F(K(x)) \arrow[ur, "\varphi_{K(x)}"'] &
\end{tikzcd}$$

Note that, given the first two commutative squares, we can stack them to create a larger commutative square:

$$
\begin{tikzcd}
F(x) \arrow[r] \arrow[d, "\varphi_x"'] \arrow[dd, bend right=60, "F(\eta_{K(\gamma(x))})"'] & F(y) \arrow[d, "\varphi_y"] \arrow[dd, bend left=60, "F(\eta_{K(\gamma(y))})"] \\
G(\Gamma(x)) \arrow[r] \arrow[d, "\psi_{\Gamma(x)}"] & G(\Gamma(y)) \arrow[d, "\psi_{\Gamma(y)}"'] \\
F(K(\Gamma(x))) \arrow[r] & F(K(\Gamma(y)))
\end{tikzcd}
\qquad
\begin{tikzcd}
G(x) \arrow[r] \arrow[d, "\psi_x"'] \arrow[dd, bend right=60, "G(\eta_{\Gamma(K(x))})"] & G(y) \arrow[d, "\psi_y"] \arrow[dd, bend left=60, "G(\eta_{\Gamma(K(y))})"] \\
F(K(x)) \arrow[r] \arrow[d, "\varphi_{K(x)}"'] & F(K(y)) \arrow[d, "\varphi_{K(y)}"'] \\
G(\Gamma(K(x))) \arrow[r] & G(\Gamma(K(y)))
\end{tikzcd}
$$

If the two triangular diagrams did not hold, then we would we would see that there would be two different, but not necessarily equal ways of getting from $F$ to $F(K(\Gamma))$ and $G$ to $G(\Gamma(K(x)))$. Note also that, if we really wanted to, we could keep stacking these diagrams on and on.

The interleaving of two functors satisfies the following three properties.

**Proposition 11.2.4** (Functoriality)**.** Let $\Gamma, K$ be translations on a preordered set $P$. If $F, G \in \mathcal{D}^P$, and if $F, G$ are $(\Gamma, K)$-interleaved, then $H \circ F$ and $H \circ G$ are also $(\Gamma, K)$ interleaved.

*Proof.* This is true since any functor applied to a commutative diagram will output a commutative diagram. Thus if we compose $H$ with the commutative diagrams which arise from the interleaving of $F, G$, we get

$$
\begin{tikzcd}
H \circ F(x) \arrow[r] \arrow[d, "H(\varphi_x)"'] & H \circ F(y) \arrow[d, "H(\varphi_y)"] \\
H \circ G(\Gamma(x)) \arrow[r] & H \circ G(\Gamma(y))
\end{tikzcd}
\qquad
\begin{tikzcd}
H \circ G(x) \arrow[r] \arrow[d, "H(\psi_x)"'] & H \circ G(y) \arrow[d, "H(\psi_y)"] \\
H \circ F(K(x)) \arrow[r] & H \circ F(K(y))
\end{tikzcd}
$$

$$
\begin{tikzcd}
H \circ F(x) \arrow[rr, "H(F(\eta_{K(\Gamma(x))}))"] \arrow[dr, "H(\varphi_x)"'] & & H \circ F(K(\Gamma(x))) \\
& H \circ G(\Gamma(x)) \arrow[ur, "H(\psi_{\Gamma(x)})"'] &
\end{tikzcd}
$$

$$
\begin{tikzcd}
H \circ G(x) \arrow[rr, "H(G(\eta_{\Gamma(K(x))}))"] \arrow[dr, "H(\psi_x)"'] & & H \circ G(\Gamma(K(x))) \\
& H \circ F(K(x)) \arrow[ur, "H(\varphi_{K(x)})"'] &
\end{tikzcd}
$$

The above diagrams can be reconciled with the definition of an $(\Gamma, K)$ interleaving, so that $H \circ F, H \circ G$ are $(\Gamma, K)$ are interleaved. ∎

**Proposition 11.2.5** (Monotonicity)**.** Let $\Gamma_1, \Gamma_2, K_1, K_2$ be translations of a preordered set $P$ such that $\Gamma_1 \leq \Gamma_2$ and $K_1 \leq K_2$. If two persistence modules $F, G \in \mathcal{D}^P$ are $(\Gamma_1, K_1)$ interleaved, then they are also $(\Gamma_2, K_2)$ interleaved.

*Proof.* Since $\Gamma_1 \leq \Gamma_2$ and $K_1 \leq K_2$, there must exist natural transformations $\alpha : \Gamma_1 \longrightarrow \Gamma_2$ and $\beta : K_1 \longrightarrow K_2$. Now since $F, G$ are $(\Gamma_1, K_1)$-interleaved, this means we get the usual diagrams, but we can stack an extra layer on the bottom.

$$
\begin{tikzcd}
F(x) \arrow[r] \arrow[d, "\varphi_x"'] & F(y) \arrow[d, "\varphi_y"] \\
G(\Gamma_1(x)) \arrow[r] \arrow[d, "G(\alpha_x)"'] & G(\Gamma_1(y)) \arrow[d, "G(\alpha_y)"] \\
G(\Gamma_2(x)) \arrow[r] & G(\Gamma_2(y))
\end{tikzcd}
\qquad
\begin{tikzcd}
G(x) \arrow[r] \arrow[d, "\psi_x"'] & G(y) \arrow[d, "\psi_y"] \\
F(K_1(x)) \arrow[r] \arrow[d, "F(\beta_x)"'] & F(K_1(y)) \arrow[d, "F(\beta_y)"] \\
F(K_2(x)) \arrow[r] & F(K_2(y))
\end{tikzcd}
$$

Hence we can see our natural transformations of interest are $G(\alpha) \circ \varphi : F \longrightarrow G \circ \Gamma_2$ and $F(\beta) \circ \psi : G \longrightarrow F \circ K_2$. We now have to show that our two required triangular diagrams must commute. Towards this goal, consider the diagram below.

$$
\begin{tikzcd}
F(x) \arrow[r, "F(\eta_{K_1(\Gamma_1(x))})"] \arrow[dr, "\varphi_x"'] & F(K_1(\Gamma_1(x))) \arrow[r, "F(\beta_{\Gamma_1(x)})"] & F(K_2(\Gamma_1(x))) \arrow[r, "F(K_2(\alpha_x))"] & F(K_2(\Gamma_2(x))) \\
& G(\Gamma_1(x)) \arrow[u, "\psi_{\Gamma(x)}"'] \arrow[r, "G(\alpha_x)"'] & G(\Gamma_2(x)) \arrow[ur, "F(\beta_x) \circ \psi_x"'] &
\end{tikzcd}
$$

The left triangle commutes since $F, G$ are a $(\Gamma_1, K_1)$ interleaving, while the rightmost commutes by the original square diagrams. We've outlined their correspondence in colors. We almost have what we want, but we need to make sure $F(K_2(\alpha_x)) \circ F(\beta_{\Gamma_1(x)}) \circ F(\eta_{\Gamma_1(K_1(x))}) = F(\eta_{\Gamma_2(K_2(x))})$. To do this, observe that the diagram

$$
\begin{tikzcd}
x \arrow[r, "\eta_{K_2\Gamma_2}"] \arrow[d, "\eta_{K_1\Gamma_1}"'] & K_2(\Gamma_2(x)) \\
K_1(\Gamma_1(x)) \arrow[r, "\beta_{\Gamma_1(x)}"'] & K_2(\Gamma_1(x)) \arrow[u, "K_2(\alpha_x)"']
\end{tikzcd}
$$

must necessarily commute as it is a diagram inside of $P$, a preordered set. Therefore, the image of this diagram under $F$ must produce a commutative diagram, so that we do in fact get our desired relation. All together, we then have

$$\begin{tikzcd}
F(x) \arrow[rr, "F(\eta_{K_2\Gamma_1})"] \arrow[dr, "(G(\alpha)\circ\psi)_x"'] & & F(K_2(\Gamma_1(x))) \\
& G(\Gamma_2(x)) \arrow[ur, "(F(\beta)\circ\psi)_x"'] &
\end{tikzcd}$$

The same procedure can be repeated dually to demonstrate commutativity for the other required triangular diagram. Thus we have that $F, G$ are $(\Gamma_2, K_2)$-interleaved. ■

**Proposition 11.2.6** (Triangle inequality.)**.** Let $\Gamma_1, \Gamma_2, K_1, K_2$ be translations of a preordered set $P$. Suppose $F, G, H \in \mathcal{D}^P$. Then if $F, G$ are $(\Gamma_1, K_1)$-interleaved and $G, H$ are $(\Gamma_2, K_2)$-interleaved, then $F, H$ are $(\Gamma_2 \circ \Gamma_1, K_1 \circ K_2)$-interleaved.

*Proof.* First observe that since $F, G$ are $(\Gamma_1, K_1)$-interleaved and $G, H$ are $(\Gamma_2, K_2)$-interleaved, we have the natural transformations

$$\varphi : F \longrightarrow G \circ \Gamma_1 \qquad\qquad \varphi' : G \longrightarrow H \circ \Gamma_2$$
$$\psi : G \longrightarrow F \circ K_1 \qquad\qquad \psi' : H \longrightarrow G \circ K_2$$

which satisfy the required diagrams. Consider the diagrams

$$\begin{tikzcd}
F(x) \arrow[r] \arrow[d, "\varphi_x"'] & F(y) \arrow[d, "\varphi_y"] \\
G(\Gamma_1(x)) \arrow[r] \arrow[d, "\varphi'_{\Gamma_1(x)}"'] & G(\Gamma_1(y)) \arrow[d, "\varphi'_{\Gamma_1(y)}"] \\
H(\Gamma_2(\Gamma_1(x))) \arrow[r] & H(\Gamma_2(\Gamma_1(y)))
\end{tikzcd}
\qquad
\begin{tikzcd}
H(x) \arrow[r] \arrow[d, "\psi'_x"'] & H(y) \arrow[d, "\psi'_y"] \\
G(K_2(x)) \arrow[r] \arrow[d, "\psi_{K_2(x)}"'] & G(K_2(y)) \arrow[d, "\varphi'_{K_2(y)}"] \\
F(K_1(K_2(x))) \arrow[r] & F(K_1(K_2(y)))
\end{tikzcd}$$

which commute by our given interleavings. Then there are natural transformations $\psi'_{\Gamma_1} \circ \varphi : F \longrightarrow H(\Gamma_2 \circ \Gamma_1)$ and $\varphi'_{K_2} \circ \psi : H \longrightarrow F(K_1 \circ K_2)$. We now must check they satisfy the required triangular diagrams. We can demonstrate this for at least one; Consider the diagram

$$\begin{tikzcd}
F(x) \arrow[rr, "F(\eta_{K_1(\Gamma_1(x))})"] \arrow[dr, "\varphi_x"'] & & F(K_1(\Gamma_1(x))) \arrow[rr, "F(K_1(\eta_{K_2\Gamma_2})\Gamma_1)"] & & F(K_1(K_2(\Gamma_2(\Gamma_1(x))))) \\
& G(\Gamma_1(x)) \arrow[ur, "\psi_{\Gamma_1(x)}"'] \arrow[rrr, "G(\eta_{K_2\Gamma_2}(\Gamma_1(x)))"] \arrow[dr, "\varphi'_{\Gamma_1(x)}"'] & & & G(K_2(\Gamma_2(\Gamma_1(x)))) \arrow[u, "\psi_{K_2(\Gamma_2(\Gamma_1(x)))}"'] \\
& & H(\Gamma_2(\Gamma_1(x))) \arrow[urr, "\psi'_{\Gamma_2(\Gamma_1(x))}"'] & &
\end{tikzcd}$$

The above diagram commutes by our given interleavings. The diagram in blue commutes since

$F, G$ are $(\Gamma_1, K_1)$ interleaved, while the diagram in red commutes since $G, H$ are $(\Gamma_2, K_2)$-interleaved.

∎

## 11.3 Interleaving Distances via Sublinear Projections and Superlinear Families

**Definition 11.3.1.** A **sublinear projection** is a function $\omega : \mathbf{Trans}_P \longrightarrow [0, \infty]$ which acts on the objects of $\mathbf{Trans}_P$ in such a way that $\omega_I = 0$ and $\omega_{\Gamma_1\Gamma_2} \leq \omega_{\Gamma_1} + \omega\Gamma_2$.

Moreover, we say a sublinear projection is **monotone** if whenever $\Gamma \leq K$ we have that $\omega_\Gamma \leq \omega_K$.

Note that we can turn a sublinear projection $\omega$ into a monotone one by defining

$$\overline{\omega}_\Gamma = \inf\{\omega_{\Gamma'} \mid \Gamma' \geq \Gamma\}.$$

This is monotone since, if $\Gamma \leq K$ is a pair of translations, then one can observe that

$$\{\omega_{\Gamma'} \mid \Gamma' \geq \Gamma\} \supset \{\omega_{\Gamma'} \mid \Gamma' \geq K\} \implies \overline{\omega}_\Gamma \leq \overline{\omega}_K.$$

Also note another nice property: for every sublinear projection $\omega$, it is always the case that $\overline{\omega}_\Gamma \leq \omega_\Gamma$ for any translation $\Gamma$.

**Definition 11.3.2.** Suppose $F, G$ are interleaved by a pair of translations $(\Gamma, K)$. Then we say $F, G$ are **$\varepsilon$-interleaved** with respect to $\omega$ if

$$\omega_\Gamma, \omega_K \leq \varepsilon.$$

Now we prove a small lemma.

**Lemma 11.3.3.** Let $\omega$ be a sublinear projection on a preorder $P$, and let $\Gamma$ be a translation of $P$. Then for every $\eta > 0$, there exists a translation $\Gamma' \geq \Gamma$ such that

$$\omega_{\Gamma'} \leq \overline{\omega}_\Gamma + \eta.$$

*Proof.* Suppose the statement was false. Then this would imply the existence of some $\eta > 0$ with the property that

$$\overline{\omega}_\Gamma + \eta < \omega_{\Gamma'}$$

for all $\Gamma' \geq \Gamma$. Hence we would see that

$$\overline{\omega}_\Gamma \neq \inf\{\omega_{\Gamma'} \mid \Gamma' \geq \Gamma\}$$

which is a contradiction. ∎

With the definition of a sublinear projection, we can now create a (psuedo)metric between persistence modules.

**Definition 11.3.4.** Let $F, G \in \mathcal{D}^P$, and suppose $\omega$ is a sublinear projection. Then their interleaving distance is given by

$$\begin{aligned} d^\omega(F,G) &= \{\varepsilon \in [0,\infty) \mid F,G \text{ are } \varepsilon\text{-interleaved w.r.t. } \omega\} \\ &= \{\varepsilon \in [0,\infty) \mid F,G \text{ are } (\Gamma,K)\text{-interleaved and } \omega_\Gamma,\omega_K \le \varepsilon\}. \end{aligned}$$

**Proposition 11.3.5.** Let $\omega$ be a sublinear projection. Then $d^\omega = d^{\overline{\omega}}$.

*Proof.* We will prove this by first showing that $d^\omega \ge d^{\overline{\omega}}$, and then demonstrating that $d^\omega - d^{\overline{\omega}} = 0$.

$\boldsymbol{d^\omega \ge d^{\overline{\omega}}}$ If a pair of persistence modules $F, G$ are $\varepsilon$-interleaved by $(\Gamma, K)$ with respect to $\omega$, then we can observe that

$$\overline{\omega}_\Gamma \le \omega_\Gamma \le \varepsilon \qquad \overline{\omega}_K \le \omega_K \le \varepsilon$$

so that $F, G$ are also $\varepsilon$-interleaved by $(\Gamma, K)$ with respect to $\overline{\omega}$. Therefore,

$$\{\varepsilon \in [0,\infty) \mid F,G \text{ are } \varepsilon\text{-interleaved w.r.t. } \omega\} \subseteq \{\varepsilon \in [0,\infty) \mid F,G \text{ are } \varepsilon\text{-interleaved w.r.t. } \overline{\omega}\}.$$

If we take the infimum of the above relation, we get that $d^{\overline{\omega}} \le d^\omega$.

$\boldsymbol{d^\omega - d^{\overline{\omega}} = 0}$**.** Let $\delta > 0$. We'll show that for any persistence modules $F, G$ that

$$d^\omega(F,G) - d^{\overline{\omega}}(F,G) \le \delta$$

which, in combination of the fact that $d^{\overline{\omega}} \le d^\omega$, will then give us our result. Towards this goal, let $\Gamma, K$ be an interleaving of $F, G$ such that

$$\overline{\omega}_\Gamma, \overline{\omega}_K \le d^{\overline{\omega}}(F,G) + \delta.$$

Such an interleaving must exist or else $d^{\overline{\omega}}(F,G)$ is larger than we thought. By the lemma we proved earlier, we know that there exist translations $\Gamma', K'$ such that

$$\Gamma \le \Gamma' \qquad K \le K'$$

and

$$\omega_{\Gamma'} \le \overline{\omega}_\Gamma \le d^\omega(F,G) + \delta \qquad \omega_{K'} \le \overline{\omega}_K \le d^\omega(F,G) + \delta.$$

Note that by Monotonocity of interleavings, since $F, G$ are interleaved by $(\Gamma, K)$, we know that $F, G$ are interleaved by $(\Gamma', K')$. Therefore, we can conclude that since $\omega_{\Gamma'}, \omega_{K'} \le d^{\overline{\omega}} + \delta$, we see that

$$d^\omega(F,G) \le d^{\overline{\omega}}(F,G) + \delta \implies d^\omega(F,G) - d^{\overline{\omega}}(F,G) \le \delta.$$

Since $\delta > 0$ was arbitrary, and because $d^\omega \geq d^{\overline{\omega}}$ we have that they must be equal, as desired.

■

We now introduce an important implication of these results.

**Theorem 11.3.6.** For any sublinear translation $\omega : \mathbf{Trans}_P \longrightarrow [0, \infty]$, The interleaving distance $d = d^\omega$ becomes an extended psuedometric on $\mathcal{D}^P$.

*Proof.* To show this, we must show that $d(F, F) = 0$ for any persistence module $F$, $d$ is symmetric, and that $d$ obeys the triangle inequality.

$\boldsymbol{d(F,F) = 0}$ Observe that $d(F, F) = 0$. This is because if we denote $I : P \longrightarrow P$ to be the identity translation on $P$, then $F$ is $(I, I)$ interleaved with itself. But recall that $\omega_I = 0$.

$\boldsymbol{d(F,G) = d(G,F)}$ Now observe that $d(F, G) = d(G, F)$. This is because of the inherent symmetry present in the definition of an interleaving, which allows us to swap $F$ and $G$.

**Triangle Inequality** Finally, we show that $d$ obeys the triangle inequality. Consider a triple of persistence modules $F, G, H$. Suppose $F, G$ are $\varepsilon$ interleaved, while $G, H$ are $\varepsilon'$ interleaved. Regardless of whether or not $\varepsilon \leq \varepsilon'$ or vice versa, we know that there exist translations $\varepsilon$-translations $(\Gamma, K)$ which interleaved $F, G$ and $\varepsilon'$-translations $(\Gamma', K')$ which interleave $G, H$. By the triangle inequality of translations, we know that this implies that $F, H$ are $(\Gamma' \circ \Gamma, K \circ K')$-interleaved

Note that by sublinearity we have that

$$\omega_{\Gamma'\Gamma} \leq \omega_{\Gamma'} + \omega_\Gamma \leq \varepsilon' + \varepsilon$$
$$\omega_{KK'} \leq \omega_K + \omega_{K'} \leq \varepsilon + \varepsilon'$$

Therefore, we see that

$$d(F, H) \leq \varepsilon' + \varepsilon.$$

Taking the infimum over $\varepsilon', \varepsilon$, we get that

$$d(F, H) \leq d(F, G) + d(G, H)$$

as desired.

■

We'll now show that this isn't the only way to invent a metric for persistence modules in their functor category.

**Definition 11.3.7.** Let $P$ be a preorder. A **superlinear family** $\Omega : [0, \infty) \longrightarrow \mathbf{Trans}_P$ is a function where

$$\varepsilon \mapsto \Omega_\varepsilon \in \mathbf{Trans}_P$$

such that $\Omega_{\varepsilon_1}\Omega_{\varepsilon_2} \le \Omega_{\varepsilon_1+\varepsilon_2}$.

Note that in $\mathbf{Trans}_P$, the identity $I : P \longrightarrow P$ is an initial object. So if $\varepsilon_1 \le \varepsilon_2$, we know that

$$I \le \Omega_{\varepsilon_2-\varepsilon_1}.$$

Appending $\Omega_{\varepsilon_1}$ on the right, we get that

$$I\Omega_{\varepsilon_1} \le \Omega_{\varepsilon_2-\varepsilon_1}.$$

Using the fact that $\Omega_{\varepsilon_1}\Omega_{\varepsilon_2} \le \Omega_{\varepsilon_1+\varepsilon_2}$, we see that

$$I\Omega_{\varepsilon_1} \le \Omega_{\varepsilon_2-\varepsilon_1} \le \Omega_{\varepsilon_2}.$$

Since $I$ is the identity, we know that $I\Omega_{\varepsilon_1} = \Omega_1$. We thus have that

$$\Omega_{\varepsilon_1} \le \Omega_{\varepsilon_2}$$

so that **superlinear families are monotonic**.

Now, how does this turn into a metric?

**Definition 11.3.8.** Let $P$ be a preorder and $\mathcal{D}$ a category. Then for $F, G \in \mathcal{D}^P$, we define their **interleaving distance**

$$d^{\Omega}(F,G) = \inf\{\varepsilon \in [0,\infty) \mid F, G \text{ are } \Omega_\varepsilon\text{-interleaved}\}.$$

If the above set is empty, we set $d^{\Omega}(F,G) = \infty$.

**Theorem 11.3.9.** The interleaving distance $d^{\Omega}$ is an extended pseudometric.

*Proof.* To show this, we need to prove that for persistence modules $F, G$, $d(F,F) = 0$, $d(F,G) = d(G,F) = 0$, and that the metric satisfies the triangle inequality.

$\boldsymbol{d(F,F) = 0}$**.** Observe that the functors $F, F$ are $(I, I)$-interleaved. Given that $I \le \Omega_0$ since it is initial, we see that $d(F,F) = 0$.

$\boldsymbol{d(F,G) = d(G,F)}$**.** Observe that the definition is purely symmetric so that this result is instant.

**Triangle inequality.** Let $F, G, H$ be persistence modules and suppose $F, G$ are $\Omega_{\varepsilon_1}$-interleaved while $G, H$ are $\Omega_{\varepsilon_2}$-interleaved. Then by the triangle property of translations, we know that $F, H$ are $(\Omega_{\varepsilon_2}\Omega_{\varepsilon_1}, \Omega_{\varepsilon_1}\Omega_{\varepsilon_2})$-interleaved.

Observe that

$$\Omega_{\varepsilon_2}\Omega_{\varepsilon_1} \le \Omega_{\varepsilon_1+\varepsilon_2}$$
$$\Omega_{\varepsilon_1}\Omega_{\varepsilon_2} \le \Omega_{\varepsilon_1+\varepsilon_2}.$$

By monotonicty of translations, this implies that $F, H$ are $\Omega_{\varepsilon_1+\varepsilon_2}$-interleaved, so that

$$d^\Omega(F, H) \leq \varepsilon_1 + \varepsilon_2.$$

Taking the infimum over $\varepsilon_1, \varepsilon_2$, we get that

$$d^\Omega(F, H) \leq d^\Omega(F, G) + d^\Omega(G, H)$$

as desired.

■

## 11.4 General Persistence Diagrams

Persistence diagrams (and barcodes) give a visual representation of how a filtration of a topological space (usually a simplicial complex) evolves. It keeps track of homological dimensions which are "born" and "killed" throughout this evolution.

Let $X$ be a topological space. We know from algebraic topology that there exists a $n$-th singular homology group

$$H_n(X).$$

Suppose that $f : X \longrightarrow \mathbb{R}$ is a real-valued function. An example of this is the height function of a sphere centered at the origin. Now one thing we can do with these types of functions is take any $a \in \mathbb{R}$ and consider

$$f^{-1}((\infty, a]) \subseteq X.$$

The space $f^{-1}((\infty, a]) \subseteq X$ is a topological space induced by the subspace topology of $X$. In general, this process can be modeled functorially. Let $\mathbb{R}$ be a category with morphisms given by poset structure. Then

$$\begin{aligned} E : \mathbb{R} &\longrightarrow \textbf{Top} \\ a &\longmapsto f^{-1}((\infty, a]) \end{aligned}$$

since if $a \le b$ then this induces a continuous function

$$i : f^{-1}((\infty, a]) \longrightarrow f^{-1}((\infty, b])$$

namely, the inclusion function. We denote the functor as $E$ for "evolution," as this functor filters the space $X$. As we send $a$ to infinity, we ultimately obtain the entire topological space.

Switching focus, consider the homology group of this subspace

$$H_n(f^{-1}((\infty, a])).$$

We can *also* outline this behavior as functorial where we send

$$\begin{aligned} H : \textbf{Top} &\longrightarrow \textbf{Ab} \\ f^{-1}((\infty, a]) &\longmapsto H(f^{-1}((\infty, a])) \end{aligned}$$

since for any $a \le b$, we have a group homomorphism which we denote as $\varphi_a^b$:

$$\varphi_a^b : H(f^{-1}(\infty, a]) \longrightarrow H(f^{-1}(\infty, b]).$$

Now we can outline this overall data pipeline as a functor $H \circ E : \mathbb{R} \longrightarrow \mathbf{Ab}$

$$\begin{aligned} H \circ E : \mathbb{R} &\longrightarrow \mathbf{Top} \longrightarrow \mathbf{Ab} \\ a &\longmapsto f^{-1}((\infty, a]) \longmapsto H(f^{-1}((\infty, a])). \end{aligned}$$

What's really happening here? First, $E$ records the evolution of the topological space under $f : X \longrightarrow \mathbb{R}$. Then $H$ records the homology groups; overall, $H \circ E$ records the topological evolution! We are thus interested in the following objects.

**Definition 11.4.1.** Let $a \leq b$. Recall that

$$H \circ E(a \leq b) = \varphi_a^b.$$

Since we are interested in the *image* of these mappings, which will be a group, we denote

$$F([a, b]) = \operatorname{Im}(\varphi_a^b) = \operatorname{Im}\Big(H(f^{-1}((\infty, a])) \longrightarrow H(f^{-1}((\infty, b]))\Big)$$

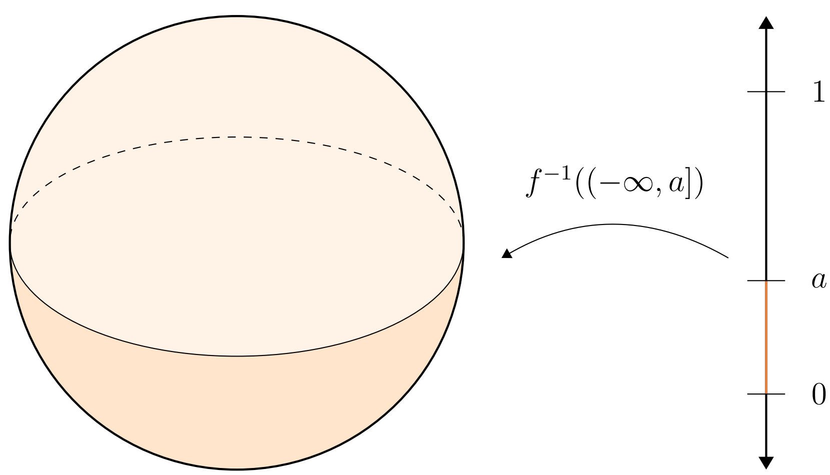
to be a **persistence homology group** from $a$ to $b$.

**Definition 11.4.2.** For a persistence homology group $F([a, b])$, define the **Betti number** from $a$ to $b$ as

$$\beta_a^b = \operatorname{rank}(F([a, b])).$$

In most nice topological spaces, the homology doesn't change much through its evolution. That is, as we move from $a$ to $b$, the persistence homology groups $F_a^b$ don't change much.

For example, if $f : X \longrightarrow \mathbb{R}$ is the height function and $X$ is a sphere, the topology will not change until we get from one pole to the other.




What does it mean for the topology to change in this context? It means that we were at some value $a$, but then at $a + \varepsilon$ the homology became different. This means that

$$H(f^{-1}((\infty, a])) \longrightarrow H(f^{-1})(\infty, a + \varepsilon]$$

is *not* an isomorphism. Finding out when the homology does change is valuable information, so we keep track of these points.

**Definition 11.4.3.** A **critical value** of $f : X \longrightarrow \mathbb{R}$ is an $a \in \mathbb{R}$ such that there exists an $\varepsilon > 0$ such that

$$H_n(f^{-1}((\infty, a-\varepsilon])) \longrightarrow H_n(f^{-1}((\infty, a+\varepsilon]))$$

is *not* an isomorphism. The function $f$ is called **tame** if $f$ has finitely many critical values.

Let $f : X \longrightarrow \mathbb{R}$ be a tame function. Then we have finitely many critical values $\{s_1, s_2, \dots, s_n\}$. Let $\{t_0, t_1, \dots, t_n\}$ be any interleaved sequence of numbers such that $t_{i-1} < s_i < t_i$. We will see soon why such a choice has much freedom in it. Now append to this sequence $t_{-1} = s_0 = -\infty$ an $t_{n+1} = s_{n+1} = \infty$.

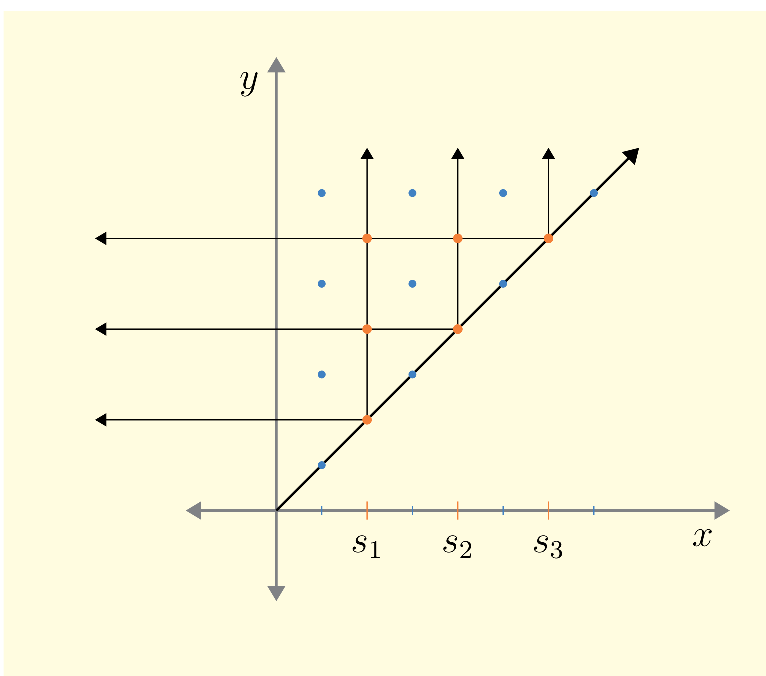
We are now ready to define persistence diagrams.

**Definition 11.4.4.** Let $f : X \longrightarrow \mathbb{R}$ be tame and $(s_i, s_j)$ be a tuple of critical values. Then we define the **multiplicity** of $(s_i, s_j)$ to be

$$\mu_i^j = \beta_{t_{i-1}}^{t_i} - \beta_{b_i}^{b_j} + \beta_{b_i}^{b_{j-1}} - \beta_{b_i}^{b_j}$$

**Definition 11.4.5.** The persistence diagram of the tame function $f : X \longrightarrow \mathbb{R}$ $D(f)$ is the *multiset* of tuples $(s_i, s_j)$ each with multiplicity $\mu_i^j$. Alternatively,

$$D(f) = \bigcup_{i=0}^{n+1} \bigcup_{j=0}^{n+1} \left( \bigcup_{k=1}^{\mu_i^j} \{(s_i, s_j)\} \right)$$



Persistence diagrams consist of points in $\mathbb{R} \times \mathbb{R} \cup \{\infty\}$ above the diagonal $y = x$. Thus let **Dgm** be the category of half open intervals $[p, q)$ with $p < q$ and intervals of the form $[p, \infty)$.

In what follows, let $S = \{s_1, s_2, \ldots, s_n\}$ be a finite set of real numbers, and let $(G, +)$ be an abelian group with identity $e$.

**Definition 11.4.6.** A map $X : \mathbf{Dgm} \longrightarrow G$ is $S$**-constructible** if for every $I \subseteq J$ where

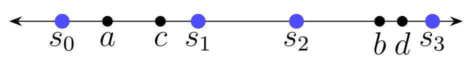
$$J \cap S = I \cap S$$

we have $X(I) = X(J)$.

The motivation for defining this type of function arises from the rank function

$$\begin{aligned}\beta_a^b &: \mathbf{Dgm} \longrightarrow \mathbb{Z}\\ &= \operatorname{rank}(F([a, b]))\\ &= \operatorname{rank}(\operatorname{Im}\left(H(f^{-1}((\infty, a])) \longrightarrow H(f^{-1}((\infty, b]))\right))\end{aligned}$$

Suppose that our critical points are $S = \{s_0, s_1, s_2, s_3\}$ and that we have two intervals $I = [a, b]$ and $J = [c, d]$ such that $I \subseteq J$ and $I \cap S = J \cap S$.




Clearly in this case we have that $I \cap S = J \cap S$. Now observe that

$$\beta_a^b = \beta_c^d$$

since these intervals observe the same changes in rank.

Therefore, we see that the rank function for a tame function $f : \mathbb{R} \longrightarrow X$ is $S$-constructible.

**Definition 11.4.7.** A map $Y : \mathbf{Dgm} \longrightarrow G$ is $S$**-finite** if

$$Y(I) \neq e \implies I = [s_i, s_j) \text{ or } I = [s_i, \infty)$$

Alternatively, this states that

$$I \neq [s_i, s_j) \text{ and } I \neq [s_i, \infty) \implies Y(I) = e.$$

which is probably a better way of thinking about this.

This leads to the following definition:

**Definition 11.4.8.** A **persistence diagram** is a finite map $Y : \mathbf{Dgm} \longrightarrow G$.

The motivation for this is due to the persistence diagram. Given a persistence diagram, we can extend it to a mapping

$$X : \mathbf{Dgm} \longrightarrow \mathbb{Z}$$
$$[a,b) \mapsto \beta_{a_1}^{b_1} - \beta_{a_2}^{b_2} + \beta_{a_2}^{b_1} - \beta_{a_1}^{b_1}$$

where $a_1 \le a \le a_2$ and $b_1 \le b \le b_2$ are values within some sufficiently small neighborhood of $a$ and $b$. Note that in this extension, if $[a,b) \ne [s_i, s_j)$ or $[s_i, \infty)$ in, then each $\beta_{a_i}^{b_j}$ is of full rank, so that

$$X([a,b)) = 0.$$

Hence we see that the persistence diagram is $S$-finite where $S$ is the finite set of critical values.

We now want to invent a distance between persistence diagrams. To do so, we must first denote $G$ as not only an abelian group, but one with a translational invariant partial ordering $\le$. What we mean by that is if $a \le b$ then $a + c \le b + c$ for any $a, b, c \in G$.

**Definition 11.4.9.** Consider $Y_1, Y_2 : \mathbf{Dgm} \longrightarrow G$ be a pair of persistence diagrams. We say there exists a **morphism** $\varphi : Y_1 \longrightarrow Y_2$ if

$$\sum_{\substack{J \in \mathbf{Dgm} \\ I \subseteq J}} Y_1(J) \le \sum_{\substack{J \in \mathbf{Dgm} \\ I \subseteq J}} Y_2(J)$$

for all $I \in \mathbf{Dgm}$.

Note the above sums are finite.

Observe that if $\varphi : Y_1 \longrightarrow Y_2$ and $\varphi' : Y_2 \longrightarrow Y_3$, then we can define the unique morphism $\varphi' \circ \varphi : Y_1 \longrightarrow Y_3$. Therefore, this morphism relation establishes a reflexive, transitive ordering on our persistence diagrams. Thus we can consider the category of persistence diagrams $\mathbf{PDiag}(G)$ into the group $G$ where the objects are persistence diagrams $Y : \mathbf{Dgm} \longrightarrow G$ and morphisms as described above. As we stated before, these morphisms make this category into a partial ordering.

Define the mapping

$$\mathbf{Grow}_\varepsilon : \mathbf{Dgm} \longrightarrow \mathbf{Dgm}$$
$$[p,q) \mapsto [p-\varepsilon, q+\varepsilon] \text{ and } [p,\infty) \mapsto [p-\varepsilon, \infty).$$

Now consider a pair of persistence modules $Y_1, Y_2 : \mathbf{Dgm} \longrightarrow G$. Since they are persistence modules, we know by definition that they are $S_1$ and $S_2$-finite for some finite sets $S_1, S_2$. With that said, observe that $Y_1 \circ \mathbf{Grow}_\varepsilon, Y_2 \circ \mathbf{Grow}_\varepsilon : \mathbf{Dgm} \longrightarrow G$ are again persistence modules since they $S_1'$ and $S_2'$ finite, where...

Therefore, we have an endofunctor on our category of persistence modules.

$$\nabla_\varepsilon : \mathbf{PDgm}(G) \longrightarrow \mathbf{PDgm}(G)$$
$$Y_1 : \mathbf{Dgm} \longrightarrow G \mapsto Y_1 \circ \mathbf{Grow}_\varepsilon : \mathbf{Dgm} \longrightarrow G.$$

Note that for any persistence modules $Y : \mathbf{Dgm} \longrightarrow G$, we have that $\nabla_\varepsilon(Y) \longrightarrow Y$ since for any interval $Y$,

$$\sum_{\substack{J \in \mathbf{Dgm} \\ I \subseteq J}} Y(J) = Y_1 \circ \mathbf{Grow}_\varepsilon$$