

Category Theory

Tobias Fritz (tobias.fritz@uibk.ac.at)

SoSe 2025

Contents

1	Introduction	3
2	Categories, Functors, and Natural Transformations	5
2.1	Categories	5
2.2	Isomorphisms	8
2.3	Monomorphisms and injectivity	10
2.4	Epimorphisms and surjectivity	11
2.5	Opposite categories and duality	12
2.6	Functors	14
2.7	Natural transformations	18
2.8	Diagrams and commutative diagrams	23
2.9	Functor categories	25
3	Universal properties	28
3.1	Representable functors	28
3.2	The Yoneda lemma	29
3.3	Universal elements	31
4	Limits and colimits	37
4.1	Categorical products	37
4.2	Equations and equalizers	46
4.3	Limits in general	48
4.4	Colimits	55
5	Adjoint functors	57
5.1	Definition and examples	57
5.2	Adjunctions via universal properties	60
6	Miscellaneous topics	63
6.1	Equivalence of categories	63
6.2	The Yoneda embedding	63
	Index of Categories	64
	References	65

Preface

Feedback on these notes is always welcome. If you find any errors, please let me know so that I can correct them.

Other literature

Here are a few other references that you might find useful. Concerning textbooks, my personal favourites are:

- ▷ Martin Brandenburg’s book “Einführung in die Kategorientheorie” [2] looks like a beautifully written book with lots of examples (although I do not know it well).
- ▷ Emily Riehl’s book “Category Theory in Context” [8] is another insightful and comprehensive introduction to category theory, written in a somewhat more advanced style.
- ▷ Paolo Perrone’s book “Starting Category Theory” [6] is an excellent gentle introduction with emphasis on explaining the intuition behind the abstract concepts.

The first two are quite close to these notes in terms of style and content (while covering plenty of additional material). Other sources that are worth mentioning are:

- ▷ Tom Leinster’s “Basic Category Theory” [4] is a concise and well-written introduction, though takes a bit of a different approach by treating adjoint functors before limits and representability.
- ▷ Saunders Mac Lane’s “Categories for the Working Mathematician” [5] is a classic and great if one wants to learn from the one of the founders of the subject, but can be challenging for beginners.
- ▷ Steve Awodey’s “Category Theory” [1] is well-regarded and widely read, but I’m only superficially familiar with it.
- ▷ There is a nice series of video lectures by Eugenia Cheng and Simon Willerton available here:

youtube.com/TheCatsters.

Acknowledgements

Many thanks to Joshua Tim Ocker and Manuel Lantschner for sharing their corrections.

1 Introduction

Mathematics can be viewed as the study of *patterns*. These are structures that repeat themselves, with minor variations, in various contexts: in the real world, in the natural sciences, and even in the abstract world of mathematics itself. Part of the power of mathematics comes from the diversity of situations in which these patterns appear.

Category theory is a mathematical theory of certain *mathematical* patterns. It is motivated by the observation that there are many patterns that appear in many different areas of mathematics. In this sense, category theory is a kind of mathematics of mathematics. Therefore, mastering category theory will not only provide you with astonishing insights into the structure of mathematics itself, but it will also equip you with a toolbox with which you can learn new mathematics more quickly and more deeply.

To be fair, I should mention that one can distinguish two approaches to the mathematics of mathematics:

- ▷ **Mathematical logic and model theory:** This approach studies the *logical* structure of mathematical theories. For example, are there statements that can neither be proven nor disproven within a given theory?
- ▷ **Category theory:** This approach focuses on how mathematical objects interact with each other through maps, and how this interaction can be used to characterize mathematical objects and to compare mathematical theories.

These two approaches are complementary, and I would encourage you to learn about mathematical logic and model theory as well. But let us get started with category theory by looking at a simple example that illustrates its way of thinking.

What are the natural numbers $\mathbb{N}$? The straightforward naive answer is that

$$\mathbb{N} = \{0, 1, 2, 3, \dots\}.$$

But what does *this* actually mean formally? What is this symbol “3”? Even worse, what do the dots “...” stand for, and can we make sense of them without invoking the natural numbers themselves?

Within set theory, one can define the natural numbers for example like this:

$$0 = \emptyset, \quad 1 = \{0\}, \quad 2 = \{0, 1\}, \quad 3 = \{0, 1, 2\}, \quad \dots \quad (1.1)$$

Is this a good definition? I would argue that *no*, it is not: this has nothing to do with either how we *think* of the natural numbers nor with how we *use* them.

As category theorists, we would instead approach this question by asking, what is the essential *structure* of the natural numbers? Which structure do they come equipped with that is most essential and that determines how we work with them? Perhaps one would first consider $\mathbb{N}$ as structured with respect to addition, multiplication, or its usual ordering. But on second thought, there is an even more basic structure that can be used to define all of these, namely the *successor* operation

$$\begin{aligned} s : \mathbb{N} &\longrightarrow \mathbb{N}, \\ n &\longmapsto n + 1. \end{aligned}$$

Together with the number 0, the successor operation lets us construct all the other structure on $\mathbb{N}$. So let us formalize this idea.

Definition 1.1.1. A *pointed successor set* is a triple (A, p, t) consisting of a set A , an element $p \in A$, and a map $t : A \rightarrow A$.

Thus the natural numbers can be viewed as a successor set $(\mathbb{N}, 0, s)$. Is it a special successor set of some sort? Yes, indeed! There is a precise sense in which it is the progenitor of all successor sets:

Theorem 1.1.2 (Recursion Principle). *The natural numbers are the **initial** successor set: for every pointed successor set (A, p, t) , there is a unique map $f : \mathbb{N} \rightarrow A$ such that $f(0) = p$ and the diagram*

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{s} & \mathbb{N} \\ f \downarrow & & \downarrow f \\ A & \xrightarrow{t} & A \end{array}$$

commutes.

For example, we can now define $3 := s(s(s(0)))$ and the operation $3 + _ : \mathbb{N} \rightarrow \mathbb{N}$ as the unique map such that

$$3 + 0 = 3, \quad 3 + s(n) = s(3 + n).$$

You may want to convince yourself that this amounts to applying Theorem 1.1.2 to the successor set $(\mathbb{N}, 3, s)$. Also proof by induction can be viewed as essentially a special case of Theorem 1.1.2, although making this precise is best done in a framework that is related to but different from category theory, namely dependent type theory [7].

As we will see later on, we can use Theorem 1.1.2 as the *definition* of the natural numbers, because it characterizes them uniquely up to isomorphism, and an isomorphism simply amounts to a relabeling of the elements. In other words, we have now given a definition of $\mathbb{N}$ which really captures the essence of how we think of the natural numbers and how we use them, but without saying anything about what the elements of $\mathbb{N}$ actually are!

Let me make another point that is better illustrated by a slightly different example, namely the real numbers $\mathbb{R}$. What is $\mathbb{R}$? Once again there are different ways to construct it, e.g. based on Dedekind cuts or Cauchy sequences. But this time, it is less clear how to define $\mathbb{R}$ in a way that captures its essential structure, because what is the essential structure of $\mathbb{R}$? In algebra, we would think of $\mathbb{R}$ as a field, so the essential structure would be the field operations of addition and multiplication. If we are so inclined, we could also think of $\mathbb{R}$ as an *ordered field*. But in analysis, we also use the absolute value operation $|\cdot| : \mathbb{R} \rightarrow \mathbb{R}$ a lot and work with $\mathbb{R}$ as a topological space.

So which point of view is the right one? The answer is that there is no right one, because the symbol “ $\mathbb{R}$ ” actually denotes entirely different *kinds* of mathematical objects in different contexts. In algebra, $\mathbb{R}$ might just denote the field, while in analysis it might just denote the topological space (in a situation where the field operations are not needed). This means that the symbol “ $\mathbb{R}$ ” really denotes various different objects in different *categories*.

2 Categories, Functors, and Natural Transformations

2.1 Categories

A category is best thought of as a world of mathematical objects and maps between them. For example, we will have a category **Set** of sets and functions, a category **Grp** of groups and group homomorphisms, and a category **Top** of topological spaces and continuous maps. With these examples in mind, we can now give the formal definition.

Definition 2.1.1. A *category* $\mathcal{C}$ consists of

- ▷ A collection of **objects** $\text{Ob}(\mathcal{C})$,
- ▷ For each pair of objects $X, Y \in \text{Ob}(\mathcal{C})$, a collection $\mathcal{C}(X, Y)$, whose elements are called **morphisms**,
- ▷ For each triple of objects $X, Y, Z \in \text{Ob}(\mathcal{C})$, a map

$$\begin{aligned} \mathcal{C}(X, Y) \times \mathcal{C}(Y, Z) &\longrightarrow \mathcal{C}(X, Z) \\ (f, g) &\longmapsto g \circ f \end{aligned}$$

called **composition**,

- ▷ For each object $X \in \text{Ob}(\mathcal{C})$, a morphism $\text{id}_X \in \mathcal{C}(X, X)$ called **identity of** X ,

such that the following hold:

- ▷ **Associativity law:** For every quadruple of objects $W, X, Y, Z \in \text{Ob}(\mathcal{C})$ and morphisms

$$f \in \mathcal{C}(W, X), \quad g \in \mathcal{C}(X, Y), \quad h \in \mathcal{C}(Y, Z),$$

we have

$$h \circ (g \circ f) = (h \circ g) \circ f. \tag{2.1}$$

- ▷ **Identity laws:** For every two objects $X, Y \in \text{Ob}(\mathcal{C})$ and morphism $f \in \mathcal{C}(X, Y)$,

$$f \circ \text{id}_X = f, \quad \text{id}_Y \circ f = f.$$

Example 2.1.2. (a) The category **Set** has sets as objects. A morphism $f \in \text{Set}(X, Y)$ is a function $f : X \rightarrow Y$. Composition $\circ$ is given by function composition, which means that for $f \in \text{Set}(X, Y)$ and $g \in \text{Set}(Y, Z)$, the composite $g \circ f \in \text{Set}(X, Z)$ is the function defined by

$$x \mapsto g(f(x)).$$

The identity morphism $\text{id}_X \in \text{Set}(X, X)$ is the identity function on X which maps every element to itself.

The associativity law is the usual associativity of function composition. Let us recall why this holds: with objects and morphisms as in (2.1), we verify this identity by showing that it holds on every $w \in W$. This gives

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

If the identity laws are not obvious to you, then you should verify them formally as well!

- (b) The category **Grp** has groups as objects. A morphism $f \in \mathbf{Grp}(G, H)$ is defined to be a group homomorphism $f : G \rightarrow H$. Composition $\circ$ is given by ordinary function composition as well, which is well-defined since the composition of two group homomorphisms is again a group homomorphism.

We do not need to verify the associativity and identity laws for **Grp** explicitly, because they follow from the fact that group homomorphisms are functions, and we have already verified these laws for **Set**.

- (c) The category **Top** has topological spaces as objects. A morphism $f \in \mathbf{Top}(X, Y)$ is a continuous map $f : X \rightarrow Y$. Composition and identities are defined as in **Set**, and we obtain a category for the same reasons as in **Grp**.

Remark 2.1.3 (Notation). (a) It is common to write $f : X \rightarrow Y$ instead of $f \in \mathcal{C}(X, Y)$. The object X is called the **domain** or **source** of f , and Y is called the **codomain** or **target** of f .

- (b) If the target of a morphism f coincides with the source of a morphism g , which means that we have $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ for some objects X, Y, Z , then we say that f and g are **composable**. In this case, we often abbreviate $g \circ f$ by gf , similar to how in group theory one often denotes multiplication simply by juxtaposition.
- (c) We often abbreviate $X \in \mathbf{Ob}(\mathcal{C})$ by $X \in \mathcal{C}$, thinking of the objects of a category as its “elements”.
- (d) A morphism $f : X \rightarrow X$ where source and target coincide is called an **endomorphism** of X .
- (e) Sometimes we also write

$$\mathbf{Mor}(\mathcal{C}) := \bigsqcup_{X, Y \in \mathcal{C}} \mathcal{C}(X, Y)$$

for the collection of all morphisms in $\mathcal{C}$, where $\bigsqcup$ denotes the disjoint union.

Remark 2.1.4 (Size issues). You may wonder what a “collection” is in Definition 2.1.1. Why do we not say “set”? First of all, we cannot say “set”: we want **Set** to be a category, but the collection of all sets is not a set.¹

Therefore we need allow the collection of objects to be “larger” than a set. This can be formalized in various ways, in which case one often uses the word *class* to refer to a collection that may be larger than any set. However, we will not need to worry about the formal details in these notes, since relying on our intuition about what a “collection” is and how it behaves will be sufficient for our purposes. But it is worth keeping in mind that our vagueness about what a “collection” is makes Definition 2.1.1 not fully formal as it stands.

There are categories in which the objects *do* form a set. If in addition all the collections of morphisms $\mathcal{C}(X, Y)$ are sets, then we say that $\mathcal{C}$ is a **small** category. For example, let K be a field. Then there is a category **Mat** $_K$ in which the objects are natural numbers, and a morphism $f : n \rightarrow m$ is an $m \times n$ matrix with entries in K . Composing a matrix $f : n \rightarrow m$ with a matrix $g : m \rightarrow p$ is done by matrix multiplication, and the identity morphism $\text{id}_n : n \rightarrow n$ is the identity matrix. Then the associativity and identity laws are basic properties of matrix multiplication. Since there are only countably many objects, **Mat** $_K$ is a small category.

Categories that are not small are called **large**. Thus **Set**, **Grp**, and **Top** are all large categories. But they are still **locally small**, which means that for every two objects X and Y , the collection of morphisms $\mathcal{C}(X, Y)$ is small enough to be a set, which we call the **hom-set** from X to Y . In our example, the local smallness holds because the hom-set from X to Y is (a subset of) the set of functions $X \rightarrow Y$. Trivially, every small category is locally small.

¹See e.g. Russell’s paradox for why the axioms of set theory are such that the formation of a set of all sets is not possible.

The large categories that one encounters in mathematical practice tend to be locally small (with some rare exceptions).

Although we have used the idea of “mathematical universes” as a motivation for the definition of category, there also are interesting examples that are not of this nature.

Example 2.1.5. Let G be a group. Then there is a category $\mathbf{BG}$ with a single object $*$, the hom-set

$$\mathbf{BG}(*, *) = G$$

and composition given by multiplication in G .² Then $\mathbf{BG}$ is a (small) category with a single object.

More generally, we can do the same construction for any *monoid*, where a monoid is an algebraic structure defined like a group, but without the requirement that inverses must exist.

Definition 2.1.6. A **monoid** is a set M equipped with a binary operation $\cdot : M \times M \rightarrow M$ that is associative and has a neutral element $e \in M$.

For example, the natural numbers $\mathbb{N}$ form a monoid under addition with neutral element 0, and a monoid under multiplication with neutral element 1, but neither of these monoids is a group. As another example, take any object X in a locally small category $\mathcal{C}$. Then $\mathcal{C}(X, X)$ is a monoid under composition with neutral element id_X . Since (as mentioned above) morphisms $X \rightarrow X$, where source and target coincide, are called endomorphisms of X , we also call the monoid $\mathcal{C}(X, X)$ the **endomorphism monoid** of X .

So given a monoid M , we can define a category $\mathbf{BM}$ with a single object $*$ just as in Example 2.1.5. Conversely, every small category $\mathcal{C}$ with a single object is of this form: if we denote the unique object by $*$, then we have $\mathcal{C} = \mathbf{BC}(*, *)$.

Example 2.1.7. Let $(X, \leq)$ be a **preordered set**, i.e. a set X equipped with a relation $\leq$ that is transitive and reflexive:

$$x \leq y \wedge y \leq z \implies x \leq z, \quad x \leq x.$$

Here are some concrete examples:

- ▷ $\mathbb{N}$ forms a preordered set under the usual ordering.
- ▷ $\mathbb{N}$ is also under a preordered set if define $x \leq y$ to hold if and only if x divides y .
- ▷ For any set A , the *discrete preorder* on A is the relation $x \leq y$ if and only if $x = y$. The *indiscrete preorder* is the relation $x \leq y$ for all $x, y \in A$.
- ▷ In the foundations of economic theory, *preference orders* play a central role.

To sketch how this works, let X be a set of items (like house, bicycle, boat, ...). Define $x \leq y$ to hold if and only if you prefer owning x over owning y . Then it is reasonable to assume that the transitivity and reflexivity hold, so that we have a preordered set.

Now any preordered set X can be considered as a category, where:

- ▷ The objects are the elements of X .
- ▷ There is exactly one morphism $x \rightarrow y$ if $x \leq y$, and none otherwise.
- ▷ Composition is clear, because there is only one possibility to compose the unique morphism $x \rightarrow y$ with the unique morphism $y \rightarrow z$, namely to the unique morphism $x \rightarrow z$; and the fact that this is possible is thanks to the transitivity of $\leq$.

²The reason for the notation $\mathbf{BG}$ is mainly historical and has to do with its role in algebraic topology.

- ▷ Similarly, the identity morphism $\text{id}_x : x \rightarrow x$ is the unique morphism $x \rightarrow x$, which exists by the reflexivity of $\leq$.
- ▷ The associativity and identity laws are also trivially satisfied.

Thus every preordered set can be considered as a category. This category has the special property that there is at most one morphism between any two objects. Conversely, we can view any such category as a preordered set by defining $x \leq y$ if and only if there is a morphism $x \rightarrow y$.

As the previous examples have shown, both monoids and preorders can be viewed as special kinds of categories. It is helpful to think of categories as a common generalization of these two concepts.

Notation 2.1.8. *From now on, we will often omit the “ $\circ$ ” symbol for composition, writing gf instead of $g \circ f$, similar to how one multiplies group elements by juxtaposition.*

Lemma 2.1.9 (Uniqueness of identities). *If $f : X \rightarrow X$ is an endomorphism in a category $\mathcal{C}$ such that $gf = g$ for all $g : X \rightarrow X$, then $f = \text{id}_X$. In particular, the identity morphism of an object X in a category $\mathcal{C}$ is unique.*

Proof. Taking $g = \text{id}_X$ gives

$$f = \text{id}_X f = \text{id}_X,$$

where the first equality holds because id_X is the identity morphism and the second by assumption. $\square$

2.2 Isomorphisms

We now look at a few special classes of morphisms. Doing this will allow us to get a first glimpse of how the language of categories can be used to express and generalize familiar concepts that are usually formulated in terms of elements in terms of morphisms.

For any kind of mathematical structure that you have come across, you will have encountered the notion of an *isomorphism*. Sometimes these are not necessarily called isomorphisms: for example *homeomorphisms* in topology are isomorphisms of topological spaces, but one uses a different name for purely historical reasons. But the intuitive idea is the same: an isomorphism is a structure-preserving map that has an inverse structure-preserving map. This leads to the idea that two isomorphic objects are perfectly indistinguishable from the point of view of the structure in question.

Definition 2.2.1. *A morphism $f : X \rightarrow Y$ in a category $\mathcal{C}$ is an **isomorphism** if there exists an **inverse** morphism $g : Y \rightarrow X$ such that*

$$gf = \text{id}_X, \quad fg = \text{id}_Y.$$

*Two objects X and Y are **isomorphic** if there exists an isomorphism $X \rightarrow Y$.*

Of course, every identity morphism is an isomorphism that is its own inverse. But let’s look at some less trivial examples.

Example 2.2.2. In **Set**, the isomorphisms are exactly the bijections.

Example 2.2.3. Let **Mon** be the category of monoids, i.e. with monoids as objects and monoid homomorphisms as morphisms.³

³A monoid homomorphism $f : M \rightarrow N$ is a map that preserves the multiplication and the neutral element, i.e. $f(mm') = f(m)f(m')$ for all $m, m' \in M$ and $f(e_M) = e_N$. Note that this is the exact same definition as that of group homomorphism.

Here is my favourite example of an isomorphism in **Mon**. Let $(p_n)_{n \in \mathbb{N}}$ be the sequence of prime numbers

$$p_0 = 2, \quad p_1 = 3, \quad p_2 = 5, \quad \dots$$

Let further $(\mathbb{N}, +)$ the monoid of natural numbers under addition, $(\mathbb{N}, +)^{\oplus \mathbb{N}}$ the monoid of finitely supported sequences of natural numbers under addition, and $(\mathbb{N}_{>0}, \cdot)$ the monoid of positive natural numbers under multiplication. Then the map

$$\begin{aligned} (\mathbb{N}, +)^{\oplus \mathbb{N}} &\longrightarrow (\mathbb{N}_{>0}, \cdot), \\ (x_n)_{n \in \mathbb{N}} &\longmapsto \prod_{n=0}^{\infty} p_n^{x_n} \end{aligned}$$

is a homomorphism, because

$$\prod_{n=0}^{\infty} p_n^{x_n + y_n} = \left(\prod_{n=0}^{\infty} p_n^{x_n} \right) \left(\prod_{n=0}^{\infty} p_n^{y_n} \right)$$

and $\prod_{n=0}^{\infty} p_n^0 = 1$. It is bijective, because this is exactly the statement of the fundamental theorem of arithmetic! Since it is straightforward to prove that a bijective homomorphism between monoids is an isomorphism, we have shown that prime factorization can be viewed as an isomorphism of monoids.

Example 2.2.4. In **Top**, the following fact can be counterintuitive to beginners: the unit square $[0, 1] \times [0, 1]$ is homeomorphic (=isomorphic) to the unit disk

$$D^2 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}.$$

Let us return to the general theory.

Lemma 2.2.5. *If a morphism $f : X \rightarrow Y$ is an isomorphism, then its inverse is unique.*

Proof. Suppose that $g, g' : Y \rightarrow X$ are both inverses of f . Then we have

$$g = g \circ \text{id}_Y = gf g' = \text{id}_X \circ g' = g'. \quad \square$$

Therefore we can speak of *the* inverse of an isomorphism, and we will denote it by f^{-1} . Clearly f^{-1} is also an isomorphism, and $(f^{-1})^{-1} = f$. Moreover, this guarantees that if X and Y are isomorphic, then also Y and X are isomorphic per Definition 2.2.1. The terminology would be very confusing if this were not the case!

Proposition 2.2.6. *The class of isomorphisms in a category satisfies the 2-out-of-3 property. By this we mean that for every composable morphisms $f : X \rightarrow Y$ and $g : Y \rightarrow Z$, if two out of*

$$f, \quad g, \quad gf$$

are isomorphisms, then so is the third.

Proof. This is analogous to familiar arguments in group theory, so we keep it short.

If f and g are isomorphisms, then we get

$$(gf)(f^{-1}g^{-1}) = g(ff^{-1})g^{-1} = gg^{-1} = \text{id}_Z,$$

and similarly $(f^{-1}g^{-1})(gf) = \text{id}_X$. Therefore $(gf)^{-1} = f^{-1}g^{-1}$ is the inverse of gf .

Suppose now that f and gf are isomorphisms. Then also f^{-1} is an isomorphism. By the previous part, we therefore know that $(gf)f^{-1} = g$ is an isomorphism as well. The case where g and gf are assumed to be isomorphisms is similar. $\square$

2.3 Monomorphisms and injectivity

In **Set**, a map $f : X \rightarrow Y$ is **injective** if the implication

$$f(x) = f(x') \implies x = x' \quad (2.2)$$

holds for all $x, x' \in X$. Note that this amounts to “cancelling” f on the “left”. Is there an analogous notion in an arbitrary category? At first look, one might think that the answer is no, because the structure of category alone does not allow us to talk about elements of objects. But as it turns out, we also do not *need* to talk about elements!

Definition 2.3.1. A morphism $f : X \rightarrow Y$ in a category $\mathcal{C}$ is a **monomorphism** if for all $A \in \mathcal{C}$ and every $g, h : A \rightarrow X$, the implication

$$fg = fh \implies g = h \quad (2.3)$$

holds.

One might also call this property **left cancellability** of f , and illustrate it in terms of the diagram

$$A \begin{array}{c} \xrightarrow{g} \\ \xrightarrow{h} \end{array} X \xrightarrow{f} Y$$

Proposition 2.3.2. In **Set**, **Grp** and **Top**, a morphism is a monomorphism if and only if it is injective.

Proof. Suppose that $f : X \rightarrow Y$ is a morphism in one of these categories that is injective. Then given $g, h : A \rightarrow X$ with $fg = fh$ as in the statement, we have

$$f(g(a)) = f(h(a))$$

for all $a \in A$, and therefore also $g(a) = h(a)$ by the injectivity of f . This shows the desired $g = h$.

Conversely, suppose that f is a monomorphism, and let us focus on **Set** and **Top** for the moment. Then consider the singleton set

$$1 = \{*\}.$$

This set also has a unique topology, so that we can consider 1 as an object in **Set** and in **Top**.⁴ The crucial observation is now that a morphism $g : 1 \rightarrow X$ in **Set** or **Top** is just another notation for the element $g(*) \in X$. In this way, the elements of X correspond precisely to the morphisms $1 \rightarrow X$. Under this correspondence, the implication (2.2) is just the $A = 1$ special case of the implication (2.3). In particular, if f is a monomorphism in **Set** or **Top**, then we can conclude its injectivity.

Finally, what if $f : X \rightarrow Y$ is a monomorphism in **Grp**? Then we can still look at morphisms $1 \rightarrow X$, where 1 is the trivial group containing only the neutral element. But now this is no longer informative, because homomorphisms must map the neutral element to the neutral element, and so there is only one homomorphism $1 \rightarrow X$.

Is there another way to get at the elements of X ? Yes, there is: instead of homomorphisms $g : 1 \rightarrow X$, let's look at homomorphisms $g : \mathbb{Z} \rightarrow X$, where $\mathbb{Z}$ is the group of integers under addition. Then the elements of X correspond to the homomorphisms $\mathbb{Z} \rightarrow X$,

$$\begin{aligned} \text{Grp}(\mathbb{Z}, X) &\xrightarrow{\cong} X, \\ g &\longmapsto g(1). \end{aligned}$$

This is because such a homomorphism must satisfy $g(n) = g(1)^n$, and therefore is uniquely determined by $g(1)$, and every choice of $g(1) \in X$ is possible.

⁴The same applies in **Grp**, but in this case it will not be helpful.

So to get at the injectivity, we can now make the same argument as in **Set** and **Top**, but with $\mathbb{Z}$ instead of 1. Here is an illustration of this situation:

$$\mathbb{Z} \begin{array}{c} \xrightarrow{g} \\ \xrightarrow{h} \end{array} X \xrightarrow{f} Y$$

where g and h are the homomorphisms associated to two arbitrary elements $g(1), h(1) \in X$. $\square$

2.4 Epimorphisms and surjectivity

In set-theoretic terms, a map $f : X \rightarrow Y$ is **surjective** if its image is the whole codomain,

$$\forall y \in Y \exists x \in X : f(x) = y.$$

This looks much more challenging to generalize to an arbitrary category. Surprisingly, there is a well-behaved generalization of surjectivity, which is elegantly *dual* to the definition of a monomorphism.

Definition 2.4.1. A morphism $f : X \rightarrow Y$ in a category $\mathcal{C}$ is an **epimorphism** if for all $A \in \mathcal{C}$ and every $g, h : Y \rightarrow A$, the implication

$$gf = hf \implies g = h \tag{2.4}$$

holds.

So in contrast to monomorphisms, an epimorphism is distinguished by its **right cancellability**, as illustrated by the diagram

$$X \xrightarrow{f} Y \begin{array}{c} \xrightarrow{g} \\ \xrightarrow{h} \end{array} A$$

Proposition 2.4.2. In **Set**, **Grp** and **Top**, a morphism is an epimorphism if and only if it is surjective.

Partial proof. As in the proof of Proposition 2.3.2, one direction works the same in all three categories, namely to show that every surjective morphism is an epimorphism. So let $f : X \rightarrow Y$ be a surjective morphism in one of these categories, and let $g, h : Y \rightarrow A$ be morphisms such that $gf = hf$. Then we have $g(f(x)) = h(f(x))$ for all $x \in X$. By the assumed surjectivity, this means indeed that $g(y) = h(y)$ for all $y \in Y$, and therefore $g = h$.

For the other direction, let us start with the case of **Set** and consider $f : X \rightarrow Y$ an epimorphism. There, we can consider $A := \{0, 1\}$ as our “probing object”. Then the morphisms $Y \rightarrow A$ correspond to the *subsets* of Y , since a map $g : Y \rightarrow \{0, 1\}$ is uniquely determined by the subset $g^{-1}(1) \subseteq Y$, and in other words is the *indicator function* of this subset. Under this correspondence, composing $g : Y \rightarrow A$ with $f : X \rightarrow Y$ corresponds to taking the *preimage* of the corresponding subset under f . Therefore since f is an epimorphism, we can conclude that the formation of preimages of subsets

$$f^{-1}(-) : 2^Y \longrightarrow 2^X$$

is *injective*. Since the two subsets $\text{im}(f)$ and Y both have all of X as their preimage, we can conclude the desired $\text{im}(f) = Y$.

In **Top**, we can proceed in the exact same way, provided that we make $\{0, 1\}$ with the *codiscrete* topology, where only the empty set and $\{0, 1\}$ are open. This has the property that *every* map into $\{0, 1\}$ is continuous, and therefore we still have the correspondence between morphisms into $\{0, 1\}$ and subsets.

In **Grp**, we cannot reason the same way, since there is no group that would play the role of $\{0, 1\}$. We need to show that if $f : G \rightarrow H$ is a group homomorphism that is *not* surjective, then there is a third group K and homomorphisms $g, h : H \rightarrow K$ that coincide on $\text{im}(f)$ but not on all of H . Proving such a “separation theorem”⁵ requires a bit more group theory, and would be a bit off-topic here, but it can be done! $\square$

⁵Note the similarity with other separation theorems, e.g. the Hahn–Banach theorem in functional analysis.

Example 2.4.3. Let $\mathbf{Met}_{\text{cont}}$ be the category of metric spaces and continuous maps, and consider the inclusion map $\mathbb{Q} \hookrightarrow \mathbb{R}$. Then this map is an epimorphism in $\mathbf{Met}_{\text{cont}}$, but it is not surjective! The reason is that $\mathbb{Q}$ is *dense* in $\mathbb{R}$, which implies that if $g, h : \mathbb{R} \rightarrow A$ are continuous maps that agree on $\mathbb{Q}$, then they must also agree on all of $\mathbb{R}$.

Example 2.4.4. Every map between sets that is injective and surjective has an inverse. Similarly, it is a standard fact of group theory that every bijective group homomorphism is an isomorphism, i.e. that its inverse is also a group homomorphism. Therefore in $\mathbf{Set}$ and $\mathbf{Grp}$, a monomorphism that is also an epimorphism is an isomorphism!

Is this the case in $\mathbf{Top}$ as well? There, the statement would be that every bijective continuous map is a homeomorphism. But this is false! For example,

$$\begin{aligned} [0, 2\pi) &\longrightarrow S^1, \\ t &\longmapsto e^{it} \end{aligned}$$

is a bijective continuous map, but it is not a homeomorphism. Alternatively, every set X can be equipped with the **codiscrete topology**, where the only open sets are $\emptyset$ and X , and with the **discrete topology**, where every subset is open. Then the identity map $X \rightarrow X$ from the discrete topology to the codiscrete topology is a bijective continuous map, but it is not a homeomorphism if X has more than one element.

Thus the statement that every monomorphism that is also an epimorphism is an isomorphism is true in $\mathbf{Set}$ and $\mathbf{Grp}$, but not in $\mathbf{Top}$. In fact, this statement is characteristic of the kinds of categories one encounters in algebra and tends to fail in the kinds of categories one studies in analysis.

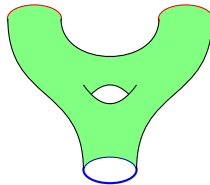
2.5 Opposite categories and duality

We have already seen a number of categories in which the morphisms are not functions. For example in $\mathbf{Mat}_K$ from Remark 2.1.4, the morphisms are matrices. Let us briefly sketch another example that is relevant in physics; this is merely for motivation and you do not need to understand it at a technical level.

Example 2.5.1 (Cobordism category, informal). For $n \in \mathbb{N}$, the **category of cobordisms** $\mathbf{Bord}_n$ has as objects oriented n -dimensional manifolds. If X and Y are such manifolds, then a **cobordism** $f : X \rightarrow Y$ is an oriented $(n+1)$ -dimensional manifold W with boundary such that the boundary is given by

$$\partial W = X \sqcup Y^{\text{op}},$$

where Y^{op} is Y with the opposite orientation. For example, here's a cobordism from two circles $S^1 \sqcup S^1$ (red) to one circle S^1 (blue):



In physical terms, we think of the incoming boundary X as a model for space at the initial time, and of the outgoing boundary Y as a model for space at the final time, while the cobordism W represents *spacetime* in between.

Now note that a cobordism $f : X \rightarrow Y$ can be composed with a cobordism $g : Y \rightarrow Z$ by sticking them together along the common boundary Y . In this way, one can show that isomorphism classes of cobordisms form a category, and it is denoted by $\mathbf{Bord}_n$. In mathematical physics, these categories appear in the definition of a *topological quantum field theory*.

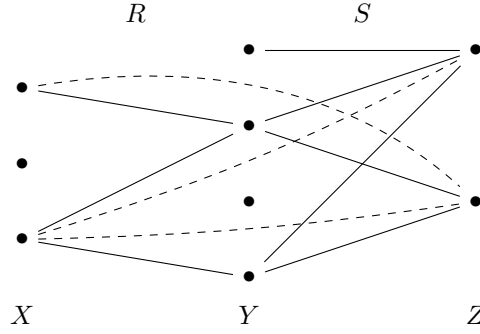


Figure 1: Two relations R and S and their composite $S \circ R$ (dashed).

Definition 2.5.2. Given a category $\mathcal{C}$, the **opposite category** $\mathcal{C}^{\text{op}}$ has the same objects as $\mathcal{C}$, but the morphisms are reversed:

$$\mathcal{C}^{\text{op}}(X, Y) := \mathcal{C}(Y, X).$$

The composition in $\mathcal{C}^{\text{op}}$ is given by the composition in $\mathcal{C}$, but with the order reversed,

$$\begin{aligned} \mathcal{C}^{\text{op}}(X, Y) \times \mathcal{C}^{\text{op}}(Y, Z) &\longrightarrow \mathcal{C}^{\text{op}}(X, Z) \\ (f, g) &\longmapsto f \circ g, \end{aligned}$$

where $\circ$ denotes composition in $\mathcal{C}$.

Note that this makes sense because in $\mathcal{C}$, we have $f : Y \rightarrow X$ and $g : Z \rightarrow Y$, so that $f \circ g : Z \rightarrow X$, which in $\mathcal{C}^{\text{op}}$ defines a morphism $X \rightarrow Z$. You may want to convince yourself that associativity holds in $\mathcal{C}^{\text{op}}$ because it holds in $\mathcal{C}$, and that $\mathcal{C}^{\text{op}}$ has the same identity morphisms as $\mathcal{C}$.

So for example in Set^{op} , a morphism $X \rightarrow Y$ is a function *from* Y *to* X . This suggests that Set^{op} behaves quite differently from Set itself. But there also are categories that look very similar to their opposite. For example in the cobordism category Bord_n from Example 2.5.1, the morphisms are hardly inherently directed, and one can always view a cobordism $f : X \rightarrow Y$ as a cobordism going the other way.⁶ Based on this, it turns out that $\text{Bord}_n^{\text{op}}$ is isomorphic as a category to Bord_n . The following examples display similar behaviour.

Example 2.5.3. Rel is the category in which objects are sets, and a morphism $X \rightarrow Y$ is a **relation** between X and Y , i.e. a subset of $X \times Y$. In other words, we have

$$\text{Rel}(X, Y) = 2^{X \times Y}.$$

You may think of a relation as (the graph of) a multivalued function: every element of x can be related to any number of elements of Y .

Given a relation $R : X \rightarrow Y$ and a relation $S : Y \rightarrow Z$, their composite $S \circ R : X \rightarrow Z$ is defined by

$$S \circ R = \{(x, z) \in X \times Z \mid \exists y \in Y : (x, y) \in R \wedge (y, z) \in S\}.$$

Figure 1 shows how this can be illustrated in terms of “paths” in a graph. You may now want to convince yourself that this composition is associative, and that for every set X , the relation

$$\text{id}_X := \{(x, x) \mid x \in X\}$$

⁶Technically one needs to be careful with orientations, as viewing f the other way will result in a cobordism $Y^{\text{op}} \rightarrow X^{\text{op}}$, but this is a minor technicality.

serves as the identity morphism.

The composition of relations implicitly appears in mathematics more frequently than one might think. For example, the definition of preordered set from Example 2.1.7 makes $\leq$ into a relation $X \rightarrow X$, and the reflexivity and transitivity axioms can now be expressed like this:

$$\text{id}_X \subseteq \leq, \quad \leq \circ \leq \subseteq \leq.$$

Now let us consider the opposite category Rel^{op} . Then a morphism $X \rightarrow Y$ in Rel^{op} is a relation $Y \rightarrow X$, i.e. a subset of $Y \times X$. But of course this is the same thing as a relation $X \rightarrow Y$, and also the composition in Rel^{op} is the same as in Rel . Thus we can say that

$$\text{Rel}^{\text{op}} = \text{Rel}.$$

Example 2.5.4. Hilb is the category in which objects are Hilbert spaces, and a morphism $X \rightarrow Y$ is a bounded linear operator $X \rightarrow Y$. Composition is composition of maps.

Also this category is self-dual: for every morphism $f : X \rightarrow Y$, we have its *adjoint* $f^* : Y \rightarrow X$ (often denoted $f^\dagger$ in physics), and this operation reverses composition,

$$(gf)^* = f^*g^*.$$

Therefore the map $f \mapsto f^*$ defines an isomorphism of categories $\text{Hilb}^{\text{op}} \cong \text{Hilb}$.

In projective geometry, a powerful idea is that of *duality*: any true statement about points and lines can be turned into a true statement about lines and points. A similar idea is at work in category theory.

Slogan 2.5.5 (Duality principle). *In category theory, every definition and every theorem that apply to all categories $\mathcal{C}$ can also be applied to the opposite category $\mathcal{C}^{\text{op}}$, where it takes the exact same form but with the directions of all arrows reversed.*

For example, the concept of monomorphism (Definition 2.3.1) is dual to that of epimorphism (Definition 2.4.1). The concept of isomorphism (Definition 2.2.1) is dual to itself.

2.6 Functors

We have seen that every type of mathematical object, like groups or topological spaces, naturally defines a category. But (small) categories are mathematical objects themselves. So, is there a *category of (small) categories*? This first of all requires having a concept of morphism between categories.

Definition 2.6.1. *Let $\mathcal{C}$ and $\mathcal{D}$ be categories. Then a **functor** $F : \mathcal{C} \rightarrow \mathcal{D}$ consists of*

- ▷ *For each object $X \in \mathcal{C}$, an object $F(X) \in \mathcal{D}$,*
- ▷ *For each morphism $f : X \rightarrow Y$ in $\mathcal{C}$, a morphism $F(f) : F(X) \rightarrow F(Y)$ in $\mathcal{D}$,*

such that the following hold:

- ▷ *For every composable pair $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ in $\mathcal{C}$, we have*

$$F(gf) = F(g)F(f), \tag{2.5}$$

- ▷ *For every object $X \in \mathcal{C}$, we have*

$$F(\text{id}_X) = \text{id}_{F(X)}. \tag{2.6}$$

As the following examples will show, functors are ubiquitous throughout mathematics. Whenever you encounter a construction of one type of mathematical object from another, you are likely dealing with a functor, and asking whether this is the case can help you understand the situation better.

Example 2.6.2 (Forgetful functors). To start with some almost trivial examples, there is a functor $\mathbf{Top} \rightarrow \mathbf{Set}$ that assigns to every topological space its underlying set, and takes every continuous map to itself. Now the equations (2.5) and (2.6) are trivially satisfied. This functor is called a **forgetful functor**, because it “forgets” the topology of a space while keeping its set of points.

There similarly is a forgetful functor $\mathbf{Grp} \rightarrow \mathbf{Set}$ that assigns to every group its underlying set, and takes every group homomorphism to the underlying function. In fact, every kind of category whose objects are “sets with extra structure” and whose morphisms are “structure-preserving maps” has a forgetful functor to $\mathbf{Set}$.

Example 2.6.3. In the other direction, there is a functor $\mathbf{Set} \rightarrow \mathbf{Top}$ that turns every set X into the topological space X with the **discrete topology**, where every subset is open. Then every function $f : X \rightarrow Y$ is continuous, because the preimage of every open set is open, and it is again obvious that this defines a functor.

We will study the relation between this functor and the forgetful functor $\mathbf{Top} \rightarrow \mathbf{Set}$ in more detail in Section 5.

Example 2.6.4 (Power set functors). There is a functor $\mathcal{P} : \mathbf{Set} \rightarrow \mathbf{Set}$ that maps every set X to its power set $\mathcal{P}(X) := 2^X$. How should this be defined on morphisms? For a function $f : X \rightarrow Y$, a natural choice to get an induced map from subsets of X to subsets of Y is to take the image. Indeed defining

$$\begin{aligned}\mathcal{P}(f) : 2^X &\longrightarrow 2^Y \\ A &\longmapsto f(A)\end{aligned}$$

works: if we have another map $g : Y \rightarrow Z$, then (2.5) holds because

$$\mathcal{P}(gf)(A) = g(f(A)) = \mathcal{P}(g)(f(A)) = \mathcal{P}(g)(\mathcal{P}(f)(A)) = (\mathcal{P}(g) \circ \mathcal{P}(f))(A)$$

for all $A \subseteq X$, while (2.6) is obvious.

Alternatively, is there a power set functor $\hat{\mathcal{P}}$ which acts on morphisms by taking *preimages*? Well, if $f : X \rightarrow Y$ is any function, then taking preimages along f indeed defines a map

$$\begin{aligned}\hat{\mathcal{P}}(f) : 2^Y &\longrightarrow 2^X \\ B &\longmapsto f^{-1}(B),\end{aligned}$$

but this now *reverses* the direction of the arrow! But fortunately, a similar calculation as before shows that this is a functor too, namely of the form $\hat{\mathcal{P}} : \mathbf{Set}^{\text{op}} \rightarrow \mathbf{Set}$.

Functors that involve an opposite category like this are often called **contravariant functors**. Sometimes when one wants to emphasize the distinction, one also speaks of a **covariant functor** when this is not the case.

Example 2.6.5 (Dual space functor). Let $\mathbf{Vect}_K$ be the category of vector spaces and linear maps over a field K . Note that every hom-set $\mathbf{Vect}_K(V, W)$ can itself be made into a vector space by defining addition and scalar multiplication pointwise.⁷ Then assigning to every $V \in \mathbf{Vect}_K$ the **dual space**

$$V^* := \mathbf{Vect}_K(V, K)$$

⁷This kind of phenomenon is studied in *enriched category theory*, which generalizes the concept of a category by allowing the hom-sets to be objects of a different category.

becomes a functor

$$(-)^* : \mathbf{Vect}_K^{\text{op}} \rightarrow \mathbf{Vect}_K \quad (2.7)$$

if we define it on morphisms as follows: to every linear map $f : V \rightarrow W$, we assign the **dual map** $f^* : W^* \rightarrow V^*$ that is defined by

$$f^*(\varphi) := \varphi \circ f.$$

This is a contravariant functor because it reverses the direction of the arrows, which we indicate by the $\cdot^{\text{op}}$ in the notation (??).

Example 2.6.6 (Hom-functors). Let $\mathcal{C}$ be any locally small category and $A \in \mathcal{C}$ an object. Then there is a **covariant hom-functor**

$$\mathcal{C}(A, -) : \mathcal{C} \rightarrow \mathbf{Set}$$

that assigns to every object $X \in \mathcal{C}$ the set of morphisms $\mathcal{C}(A, X)$, and to every morphism $f : X \rightarrow Y$ the function which post-composes with f ,

$$\begin{aligned} \mathcal{C}(A, X) &\longrightarrow \mathcal{C}(A, Y) \\ g &\longmapsto fg. \end{aligned}$$

Then (2.5) holds because of the associativity of composition, and (2.6) holds because of the identity laws. For example, if we take $\mathcal{C} := \mathbf{Top}$ and $A := 1$ the space with only one point, then the hom-functor $\mathbf{Top}(1, -) : \mathbf{Top} \rightarrow \mathbf{Set}$ is precisely the forgetful functor from Example 2.6.2.

We similarly have a **contravariant hom-functor**

$$\mathcal{C}(-, B) : \mathcal{C}^{\text{op}} \rightarrow \mathbf{Set}$$

that assigns to every object $X \in \mathcal{C}$ the set of morphisms $\mathcal{C}(X, B)$, and to every morphism $f : X \rightarrow Y$ the function which pre-composes with f ,

$$\begin{aligned} \mathcal{C}(Y, B) &\longrightarrow \mathcal{C}(X, B) \\ g &\longmapsto gf. \end{aligned}$$

Note that with $\mathcal{C} := \mathbf{Set}$ and $B := \{0, 1\}$, the hom-functor $\mathbf{Set}(-, B) : \mathbf{Set}^{\text{op}} \rightarrow \mathbf{Set}$ is essentially⁸ the contravariant power set functor $\hat{\mathcal{P}}$ from Example 2.6.4, since we can identify every subset $S \subseteq X$ with its indicator function $1_S : X \rightarrow \{0, 1\}$, as already done in the proof of Proposition 2.4.2.

Also with $\mathcal{C} := \mathbf{Vect}_K$ and $B := K$, this more or less reproduces the dual space functor from the previous example; the only difference is that the dual space V^* is now considered as a set, not as a vector space itself.

Example 2.6.7 (Taking kernels is functorial). To consider a more sophisticated example, let $\mathbf{Vect}_K^{\rightarrow}$ be the category where objects are linear maps $f : U_1 \rightarrow U_2$, and a morphism from $f : U_1 \rightarrow U_2$ to $g : V_1 \rightarrow V_2$ is a pair of linear maps $p_1 : U_1 \rightarrow V_1$ and $p_2 : U_2 \rightarrow V_2$ such that $p_2 f = g p_1$, as illustrated by the diagram

$$\begin{array}{ccc} U_1 & \xrightarrow{f} & U_2 \\ p_1 \downarrow & & \downarrow p_2 \\ V_1 & \xrightarrow{g} & V_2 \end{array} \quad (2.8)$$

We can compose a morphism $(p_1, p_2) : f \rightarrow g$ with a morphism $(q_1, q_2) : g \rightarrow h$ by taking the pair $(q_1 p_1, q_2 p_2)$. Indeed if $p_2 f = g p_1$ and $q_2 g = h q_1$, then we get

$$q_2 p_2 f = q_2 g p_1 = h q_1 p_1.$$

⁸We will make precise what this “essentially” means in Example 2.7.3.

To understand this more intuitively, it may be helpful to draw the diagram

$$\begin{array}{ccc}
 U_1 & \xrightarrow{f} & U_2 \\
 p_1 \downarrow & & \downarrow p_2 \\
 V_1 & \xrightarrow{g} & V_2 \\
 q_1 \downarrow & & \downarrow q_2 \\
 W_1 & \xrightarrow{h} & W_2
 \end{array} \tag{2.9}$$

Since associativity and identity laws are clear, we have a category $\mathbf{Vect}_K^{\rightarrow}$.

Then there is a functor

$$\ker : \mathbf{Vect}_K^{\rightarrow} \rightarrow \mathbf{Vect}_K$$

which maps every linear map $f : U_1 \rightarrow U_2$ to its **kernel**

$$\ker(f) := \{v \in U_1 \mid f(v) = 0\},$$

which is itself a vector space, and every morphism $(p_1, p_2) : f \rightarrow g$ to the linear map

$$p_1|_{\ker(f)} : \ker(f) \rightarrow \ker(g). \tag{2.10}$$

To see that p_1 indeed takes an element of $\ker(f)$ to an element of $\ker(g)$, we calculate for $v \in \ker(f)$ that

$$g(p_1(v)) = p_2(f(v)) = p_2(0) = 0,$$

as required. The induced map (2.9) can also be illustrated by extending the diagram (2.7) to the left by including the kernels,

$$\begin{array}{ccccc}
 \ker(f) & \hookrightarrow & U_1 & \xrightarrow{f} & U_2 \\
 p_1|_{\ker(f)} \downarrow & & \downarrow p_1 & & \downarrow p_2 \\
 \ker(g) & \hookrightarrow & V_1 & \xrightarrow{g} & V_2
 \end{array}$$

While (2.6) is obvious, checking (2.5) is also straightforward once one has wrapped one's head around the definitions: in the situation of (2.8), we clearly have

$$(q_1 p_1)|_{\ker(f)} = q_1|_{\ker(g)} p_1|_{\ker(f)},$$

which is the desired equation.

Example 2.6.8 (Differentiation is a functor). Let $\mathbf{Euc}$ be the category where objects are natural numbers, and a morphism $n \rightarrow m$ is a smooth function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$.⁹ Then you will be familiar with the **Leibniz rule** for the Jacobi matrix of the composition of two functions: for $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $g : \mathbb{R}^m \rightarrow \mathbb{R}^k$, we have

$$D_x(g \circ f) = D_{f(x)}(g) D_x(f), \tag{2.11}$$

where $D_x(-)$ denotes the Jacobi matrix at a point x . Furthermore, the Jacobi matrix of the identity function is the identity matrix. So doesn't this look like (2.5) and (2.6) again?

To see how this can be turned into a functor, we need to apply some trick in order to keep track of the point at which we take the derivative. So let $\mathbf{Euc}_*$ be the category where objects are *pairs* (n, x) of a natural number n and a point $x \in \mathbb{R}^n$, and a morphism $(n, x) \rightarrow (m, y)$ is a smooth function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that $f(x) = y$. Then taking the Jacobi matrix at the specified point x defines a functor

$$D : \mathbf{Euc}_* \rightarrow \mathbf{Vect}_{\mathbb{R}},$$

since the Leibniz rule (2.10) now precisely matches the composition law (2.5).

⁹The abbreviation “ $\mathbf{Euc}$ ” hints at “Euclidean space”.

Example 2.6.9 (Algebraic varieties as functors). Consider a polynomial equation or system of equations with integer coefficients, such as:

$$x^3 + y^3 = z^3.$$

This equation has a set of solutions with $x, y, z \in \mathbb{Z}$, which is a subset of $\mathbb{Z}^3$. Similarly, the set of solutions with $x, y, z \in \mathbb{C}$ is a subset of $\mathbb{C}^3$. Furthermore, every integer solution can be considered as a complex solution, so that the former set is included in the latter.

More generally, let $\mathbf{CRng}$ be the category of commutative rings and ring homomorphisms. Then for every $R \in \mathbf{CRng}$, the set of triples $x, y, z \in R$ that satisfy the equation is a subset of R^3 . Let us denote it by $V(R)$. Moreover, if $f : R \rightarrow S$ is a ring homomorphism, then applying it to x, y and z gives a triple $f(x), f(y), f(z) \in S$ that satisfies the equation, because

$$f(x)^3 + f(y)^3 = f(x^3 + y^3) = f(z^3) = f(z)^3.$$

Therefore this defines a map $V(f) : V(R) \rightarrow V(S)$. It is straightforward to check that this defines a functor $V : \mathbf{CRng} \rightarrow \mathbf{Set}$.

Proposition 2.6.10. *Let $\mathcal{C}, \mathcal{D}$ and $\mathcal{E}$ be categories and $F : \mathcal{C} \rightarrow \mathcal{D}$ and $G : \mathcal{D} \rightarrow \mathcal{E}$ functors. Then the composite $G \circ F : \mathcal{C} \rightarrow \mathcal{E}$ is a functor too.*

Of course, what we mean by the composition $G \circ F$ is that on an object $X \in \mathcal{C}$, we have $(G \circ F)(X) = G(F(X))$, and similarly on morphisms.

Proof. The main task is to verify that the composite still satisfies (2.5) and (2.6). For the former, take $h : X \rightarrow Y$ and $k : Y \rightarrow Z$ in $\mathcal{C}$ (where we avoid the notation f and g to avoid confusion with the functors F and G). Then we have in $\mathcal{E}$ that

$$(G \circ F)(kh) = G(F(kh)) = G(F(k)F(h)) = G(F(k))G(F(h)) = (G \circ F)(k)(G \circ F)(h),$$

as was to be shown. Similarly,

$$(G \circ F)(\text{id}_X) = G(F(\text{id}_X)) = G(\text{id}_{F(X)}) = \text{id}_{G(F(X))} = \text{id}_{(G \circ F)(X)},$$

which completes the proof. $\square$

From now, we will write $\mathbf{Cat}$ for the category where objects are small categories and morphisms are functors.

2.7 Natural transformations

Perhaps surprisingly, the story does not end with functors. For any two categories, we will now define a concept of *morphism between functors*. Once again, we will see that these structures appear in many different guises throughout mathematics.

Definition 2.7.1. *For categories $\mathcal{C}$ and $\mathcal{D}$ and functors $F, G : \mathcal{C} \rightarrow \mathcal{D}$, a **natural transformation** $\alpha : F \Rightarrow G$ assigns to every object $X \in \mathcal{C}$ a morphism $\alpha_X : F(X) \rightarrow G(X)$ in $\mathcal{D}$ such that for every morphism $f : X \rightarrow Y$ in $\mathcal{C}$, we have*

$$\alpha_Y \circ F(f) = G(f) \circ \alpha_X, \tag{2.12}$$

as illustrated by the diagram

$$\begin{array}{ccc} F(X) & \xrightarrow{F(f)} & F(Y) \\ \alpha_X \downarrow & & \downarrow \alpha_Y \\ G(X) & \xrightarrow{G(f)} & G(Y) \end{array} \tag{2.13}$$

The natural transformation α is a **natural isomorphism** if every α_X is an isomorphism. The functors F and G are said to be **naturally isomorphic** if there exists a natural isomorphism between them.

The equation (2.11) is often called **naturality** and its diagrammatic representation (2.12) the **naturality square**. The individual morphisms α_X are called the **components** of the natural transformation α .

As the choice of word indicates, “naturality” captures the idea that a map that is defined for an entire class of objects does not depend on arbitrary choices. You may want to see how this idea is realized in the following examples.

Example 2.7.2 (The determinant as a natural transformation). For a commutative ring R and any $n \times n$ -matrix $A \in M_n(R)$, the **determinant** $\det_R(A) \in R$ is defined by

$$\det_R(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)}.$$

Let us verify that this is a natural transformation from the functor $M_n : \mathbf{CRng} \rightarrow \mathbf{Set}$, which assigns to every ring R the set $M_n(R)$ of $n \times n$ -matrices over R , to the forgetful functor $U : \mathbf{CRng} \rightarrow \mathbf{Set}$ that assigns to every ring its underlying set. Indeed M_n acts on a ring homomorphism $f : R \rightarrow S$ simply by applying it entrywise to a matrix.

The naturality means that for every ring homomorphism $f : R \rightarrow S$, the naturality square

$$\begin{array}{ccc} M_n(R) & \xrightarrow{M_n(f)} & M_n(S) \\ \det_R \downarrow & & \downarrow \det_S \\ U(R) & \xrightarrow{U(f)} & U(S) \end{array}$$

commutes. This amounts to the following: for every $A \in M_n(R)$, we have $\det_S(M_n(f)(A)) = U(f)(\det_R(A))$, or more concretely

$$\sum_{\sigma \in S_n} \text{sgn}(\sigma) f(A_{1,\sigma(1)}) \cdots f(A_{n,\sigma(n)}) = f \left(\sum_{\sigma \in S_n} \text{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} \right).$$

This holds simply because f is a ring homomorphism.

To summarize, we can view the determinant as a natural transformation $\det : M_n \Rightarrow U$.

Given a natural isomorphism $\alpha : F \Rightarrow G$, its inverse $\beta : G \rightarrow F$ with components $\beta_X := \alpha_X^{-1}$ is automatically natural too, because the commutativity of every naturality square

$$\begin{array}{ccc} G(X) & \xrightarrow{G(f)} & G(Y) \\ \beta_X \downarrow & & \downarrow \beta_Y \\ F(X) & \xrightarrow{F(f)} & F(Y) \end{array}$$

follows automatically: it amounts to $\alpha_Y^{-1} F(f) = G(f) \alpha_X^{-1}$, which holds because of $F(f) \alpha_X = \alpha_Y G(f)$, which is the naturality of α .

Example 2.7.3. In Example 2.6.6, we were a little vague by saying that the hom-functor $\mathbf{Set}(-, \{0, 1\}) : \mathbf{Set}^{\text{op}} \rightarrow \mathbf{Set}$ is “essentially” the contravariant power set functor $\hat{\mathcal{P}}$. We can now make this precise by showing that these two functors are naturally isomorphic.

Indeed for every set X , we have a map assigning to every subset its indicator function,

$$\begin{aligned} \alpha_X : \hat{\mathcal{P}}(X) &\longrightarrow \mathbf{Set}(X, \{0, 1\}) \\ A &\longmapsto 1_A, \end{aligned}$$

and we have already observed that this is a bijection for every X . It remains to convince ourselves of the naturality. This means that we need to take any map $f : X \rightarrow Y$ and show that the naturality square

$$\begin{array}{ccc} \hat{\mathcal{P}}(Y) & \xrightarrow{\hat{\mathcal{P}}(f)} & \hat{\mathcal{P}}(X) \\ \alpha_X \downarrow & & \downarrow \alpha_Y \\ \text{Set}(Y, \{0, 1\}) & \xrightarrow{- \circ f} & \text{Set}(X, \{0, 1\}) \end{array}$$

commutes. Starting with any subset $A \subseteq Y$ on the top left, we therefore need to check that

$$1_{f^{-1}(A)} = 1_A \circ f$$

is an equality of functions $X \rightarrow \{0, 1\}$. But this is clear: for every $x \in X$, both sides evaluate to 1 if $f(x) \in A$ and to 0 otherwise.

Furthermore, every component α_X is invertible, because it has an inverse map given by

$$\begin{aligned} \beta_X : \text{Set}(X, \{0, 1\}) &\longrightarrow \hat{\mathcal{P}}(X) \\ f &\longmapsto f^{-1}(\{1\}). \end{aligned}$$

Hence α is a natural isomorphism. In particular, it follows automatically that its inverse β with components β_X is natural as well.

Example 2.7.4 (Delta distributions). Let us consider an example from probability theory. In order to keep the discussion free from measure-theoretic technicalities, we consider discrete probability only, while noting that a generalization to measure-theoretic probability is possible.

Given a set X , a **discrete probability distribution** on X is a function $p : X \rightarrow [0, 1]$ such that $\sum_{x \in X} p(x) = 1$. As may be known from analysis, this in particular requires the support $p^{-1}((0, 1])$ to be countable, since a sum of uncountably many positive numbers is necessarily infinite. Now given such a p and a function $f : X \rightarrow Y$, we can define the **pushforward distribution** $f_*p : Y \rightarrow [0, 1]$ by

$$(f_*p)(y) := \sum_{x \in f^{-1}(y)} p(x),$$

and the normalization condition is straightforward to check,

$$\sum_{y \in Y} (f_*p)(y) = \sum_{y \in Y} \sum_{x \in f^{-1}(y)} p(x) = \sum_{x \in X} p(x) = 1,$$

where the second step uses the fact that the sets $f^{-1}(y)$ for $y \in Y$ form a partition of X . Moreover if $g : Y \rightarrow Z$ is another function, then we have

$$(gf)_*(p) = g_*(f_*p),$$

since the weight of an element $z \in Z$ on the two sides is given by

$$(gf)_*(p)(z) = \sum_{x \in (gf)^{-1}(z)} p(x) = \sum_{y \in g^{-1}(z)} \sum_{x \in f^{-1}(y)} p(x) = \sum_{y \in g^{-1}(z)} (f_*p)(y) = (g_*(f_*p))(z),$$

where the second step uses the fact that $(gf)^{-1}(z)$ is the disjoint union of the sets $f^{-1}(y)$ for $y \in g^{-1}(z)$. Since additionally $\text{id}_*p = p$ is obvious, we can conclude that

$$\begin{aligned} P : \text{Set} &\longrightarrow \text{Set} \\ X &\longmapsto PX := \left\{ p : X \rightarrow [0, 1] \mid \sum_{x \in X} p(x) = 1 \right\} \\ f &\longmapsto f_* \end{aligned}$$

is a functor. We call it the **discrete probability functor**.

For every $x \in X$, we have the **delta distribution** $\delta_x \in PX$ given by

$$\delta_x(x') := \begin{cases} 1 & \text{if } x' = x, \\ 0 & \text{otherwise.} \end{cases}$$

For our purposes, it will be useful to denote this by $\delta_X(x) \in PX$ instead, where

$$\delta_X : X \longrightarrow PX$$

assigns to every point the corresponding delta distribution. Is this a natural transformation? To check this, we need to see whether for every $f : X \rightarrow Y$, the naturality square

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \delta_X \downarrow & & \downarrow \delta_Y \\ PX & \xrightarrow{Pf=f_*} & PY \end{array}$$

commutes, which means $f_* \circ \delta_X = \delta_Y \circ f$, or equivalently

$$f_*(\delta_X(x)) = \delta_Y(f(x))$$

for every $x \in X$. To prove this, we calculate the value of the left-hand side on any $y \in Y$,

$$(f_*\delta_X(x))(y) = \sum_{x' \in f^{-1}(y)} \delta_X(x)(x') = \begin{cases} 1 & \text{if } x \in f^{-1}(y), \\ 0 & \text{otherwise.} \end{cases}$$

On the other hand, for the right-hand side we clearly get 1 if $y = f(x)$ and 0 otherwise. But since $x \in f^{-1}(y)$ if and only if $y = f(x)$, we see that both sides are equal. Therefore, we can say that the formation of delta distributions is a natural transformation

$$\delta : \text{id}_{\text{Set}} \Longrightarrow P.$$

Example 2.7.5. Consider the forgetful functor $U : \text{CRng} \rightarrow \text{Set}$ from commutative rings to sets, and the functor

$$\begin{aligned} U^{(2)} : \text{CRng} &\rightarrow \text{Set} \\ R &\longmapsto R \times R \end{aligned}$$

which sends every commutative ring R to the set $R \times R$. Then the multiplication in R is a map $m_R : R \times R \rightarrow R$. As R varies, these maps define a natural transformation

$$m : U^{(2)} \Longrightarrow U.$$

Indeed for the naturality, what we need to check is that for every ring homomorphism $f : R \rightarrow S$, the naturality square

$$\begin{array}{ccc} R \times R & \xrightarrow{f \times f} & S \times S \\ \downarrow m_R & & \downarrow m_S \\ R & \xrightarrow{f} & S \end{array}$$

commutes, which means that $f(x)f(y) = f(xy)$ for every $x, y \in R$. But this holds by the very definition of ring homomorphism!

In fact, the same holds if we use *addition* instead of multiplication: it also defines a natural transformation $U^{(2)} \Rightarrow U$. If you're ambitious and looking for a fun puzzle, try to prove that the set of natural transformations $U^{(2)} \Rightarrow U$ can be identified with the set of two-variable polynomials $\mathbb{Z}[x, y]!$

Example 2.7.6. When $\mathcal{C}$ is any category and we consider the hom-functors $\mathcal{C}(A, -)$ and $\mathcal{C}(B, -)$ associated to objects A and B , then every morphism $f : A \rightarrow B$ induces a natural transformation

$$- \circ f : \mathcal{C}(B, -) \Rightarrow \mathcal{C}(A, -).$$

The naturality square for this natural transformation at some morphism $g : X \rightarrow Y$ is

$$\begin{array}{ccc} \mathcal{C}(B, X) & \xrightarrow{g \circ -} & \mathcal{C}(B, Y) \\ \downarrow - \circ f & & \downarrow - \circ f \\ \mathcal{C}(A, X) & \xrightarrow{g \circ -} & \mathcal{C}(A, Y) \end{array}$$

and the naturality holds because of the associativity law.

Later on in Section 3.2, we will see that *every* natural transformation $\mathcal{C}(B, -) \Rightarrow \mathcal{C}(A, -)$ arises in this way.

Given *three* functors $F, G, H : \mathcal{C} \rightarrow \mathcal{D}$ and natural transformations $\alpha : F \Rightarrow G$ and $\beta : G \Rightarrow H$, we can form their **vertical composition** $\beta \circ \alpha : F \Rightarrow H$, which is a natural transformation with components

$$(\beta \circ \alpha)_X := \beta_X \circ \alpha_X.$$

To see that the naturality on an arbitrary morphism $f : X \rightarrow Y$ holds, we consider the diagram

$$\begin{array}{ccc} F(X) & \xrightarrow{F(f)} & F(Y) \\ \downarrow \alpha_X & & \downarrow \alpha_Y \\ G(X) & \xrightarrow{G(f)} & G(Y) \\ \downarrow \beta_X & & \downarrow \beta_Y \\ H(X) & \xrightarrow{H(f)} & H(Y) \end{array} \quad (2.14)$$

Indeed the overall commutativity follows by a diagram chase from the commutativity of the two individual squares,

$$(\beta \circ \alpha)_Y F(f) = \beta_Y \alpha_Y F(f) = \beta_Y G(f) \alpha_X = H(f) \beta_X \alpha_X = H(f) (\beta \circ \alpha)_X.$$

A nice way to illustrate this is to draw “arrows between arrows” like this:

$$\begin{array}{ccc} & F & \\ & \downarrow \alpha & \\ \mathcal{C} & \xrightarrow{G} & \mathcal{D} \\ & \downarrow \beta & \\ & H & \end{array}$$

Example 2.7.7. Let us see how we can understand the multiplicativity of the determinant

$$\det(AB) = \det(A) \det(B) \quad (2.15)$$

as an equation between natural transformations.

For a fixed ring R , we can view its two sides as

$$\begin{array}{ccc} M_n(R) \times M_n(R) & \xrightarrow{\text{mult in } M_n(R)} & M_n(R) \\ \downarrow \det \times \det & & \downarrow \det \\ R \times R & \xrightarrow{\text{mult in } R} & R \end{array} \quad (2.16)$$

where the upper horizontal map is matrix multiplication. The commutativity of this diagram is precisely the multiplicativity of the determinant (2.14). But now as R varies, each arrow in (2.15) defines a natural transformation $U^{(2)} \Rightarrow U$ as in Example 2.7.5, and composing the arrows corresponds to vertically composing the natural transformations.

2.8 Diagrams and commutative diagrams

We have already made use of “commutative diagrams”, mainly in the form of squares like naturality squares. But what does it mean for a diagram to commute in general? And what even is a diagram in the first place?

Definition 2.8.1. (a) A **directed graph** is a tuple (V, E, s, t) where V is a set of **vertices**, E is a set of **edges**, and $s, t : E \rightarrow V$ are the **source** and **target** functions assigning to every edge its source and target vertex.

(b) A **diagram** in a category $\mathcal{C}$ is a directed graph (V, E, s, t) together with a function $\ell_V : V \rightarrow \text{Ob}(\mathcal{C})$ labelling every vertex by an object in $\mathcal{C}$, and a function $\ell_E : E \rightarrow \text{Mor}(\mathcal{C})$ such that for every edge $e \in E$, its source and target are labelled by the source and target of the label $\ell_E(e)$, meaning

$$\ell_V(s(e)) = s(\ell_E(e)) \quad \text{and} \quad \ell_V(t(e)) = t(\ell_E(e)).$$

Remark 2.8.2. While a directed graph and therefore also a diagram can have infinitely many vertices and edges, we will only need finite diagrams for the most part; Section 4 will be an exception.

Example 2.8.3. If $f : A \rightarrow A$ and $g : A \rightarrow X$ are morphisms in $\mathcal{C}$, then we can form diagrams like these:

$$\begin{array}{ccc} A & \xrightarrow{f} & A \\ g \downarrow & & \downarrow g \\ X & \xrightarrow{\text{id}_X} & X \end{array} \quad \text{and} \quad \begin{array}{ccc} & & \\ f \circlearrowleft & A & \xrightarrow{g} X \\ & & \circlearrowright \text{id}_X \end{array} \quad (2.17)$$

We say that a diagram **commutes** or **is commutative** if for every pair of nodes and every two paths of edges from one to the other, the composition of the morphisms along the two paths is the same. For example, the first diagram in (2.16) commutes if $gf = g$, and the second if $f = \text{id}_A$. The latter is because there also is a path at A with zero steps, which by definition evaluates to the identity morphism at A , just like how an empty product of numbers evaluates to 1.

In category theory, it is often case that one deals with a diagram that is built out of several smaller subdiagrams. In this case, one can often derive the commutativity of the whole diagram from the commutativity of the subdiagrams, as we had done for (2.13) and (2.18). This is a form of **diagram chasing**, where one follows the paths and applies the relevant equations in order to “deform” paths into each other. Here’s an example from my own research:¹⁰

$$\begin{array}{ccccc} FX \otimes FY & \xrightarrow{\quad \nabla \quad} & & & F(X \otimes Y) \\ & \searrow s & & \nearrow \mu & \\ & F(FX \otimes Y) & \xrightarrow{Ft} & FF(X \otimes Y) & \\ \text{id} \otimes \alpha \downarrow & \downarrow \alpha & & \downarrow \alpha & \\ FX \otimes GY & \xrightarrow{s} G(FX \otimes Y) & \xrightarrow{Gt} & GF(X \otimes Y) & \\ \alpha \otimes \text{id} \downarrow & \downarrow G(\alpha \otimes \text{id}) & & \downarrow G\alpha & \\ & G(GX \otimes Y) & \xrightarrow{Gt} & GG(X \otimes Y) & \\ & \nearrow s & & \searrow \mu & \\ GX \otimes GY & \xrightarrow{\quad \nabla \quad} & & & G(X \otimes Y) \end{array}$$

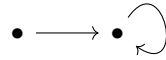
In this case, we have concluded the commutativity of the outer square from the commutativity of the inner “squares” (some of which look more like trapezoids in this drawing). One can do this by

¹⁰See the proof of Proposition C.5 in [3], which uses slightly different notation.

diagram chasing if one starts with the path given by the top and right edges, and then gradually deforms it into the path given by the bottom and left edges by using the commutativity of the inner squares. Note that diagram chasing does not always work in a “planar” way like this, and the fact that it does in this case is also because we have chosen the “right” way to draw the diagram in the plane.

It also should be mentioned that *not* all diagrams in category theory are expected to commute, and sometimes one specifies in words which subdiagrams commute and which ones do not necessarily commute.

Example 2.8.4. Consider the directed graph given by



Then a pointed successor set (in the sense of Definition 1.1.1) can be seen as a diagram in **Set** of this form where the left vertex gets mapped to a singleton set 1,

$$1 \xrightarrow{p} A \overset{t}{\curvearrowright}$$

since the data of a morphism $p : 1 \rightarrow A$ exactly amounts to choosing an element of A . This diagram is not required to commute, since we want to allow $t \neq \text{id}_A$.

A good way to think about a pointed successor set is as a **dynamical system** in discrete time: A is a set of *states* of the system, t is the *dynamics* which takes a state and transitions it in time by one step, while p selects an initial state. We may think of a general diagram of that form, where the left vertex gets mapped to an arbitrary set B instead of 1, as a dynamical system with various choices of initial conditions parametrized by B .

We can also generalize in a different way, replacing **Set** by other categories. Using **Top** instead gives us a meaningful concept of topological dynamical system, while using a suitable category of “probabilistic maps” (namely Markov kernels) gives us a category of Markov chains. If we use **Rel** instead, then we obtain a concept of *automaton* in the theoretical computer science sense.

Now suppose that we’re given two diagrams of the same shape, i.e. diagrams with the same underlying directed graph but potentially different labels. Then we can consider a **morphism of diagrams**, defined like a natural transformation: it has one component for each node of the diagram, and for each edge in the diagram the resulting naturality square must commute. In this way, we obtain a **category of diagrams**, where these morphisms compose componentwise like natural transformations.

Example 2.8.5 (Arrow category, I). Consider the graph consisting of a single edge between two vertices:



Then for any category $\mathcal{C}$, a diagram of this form is simply an arbitrary morphism f in $\mathcal{C}$, and a morphism of diagrams is a commutative square with the source and target morphisms on opposite sides. With composition being given by juxtaposition of the squares and composing on the two “long” sides, we obtain a category of diagrams called the **arrow category** $\mathcal{C}^\rightarrow$, named after the fact that the objects of $\mathcal{C}^\rightarrow$ are the “arrows” (morphisms) of $\mathcal{C}$. More explicitly, for $f : A \rightarrow B$ and $g : C \rightarrow D$ in $\mathcal{C}$, the hom-sets are given by

$$\mathcal{C}^\rightarrow(f, g) := \{(k, \ell) \in \mathcal{C}(A, C) \times \mathcal{C}(B, D) \mid gk = \ell f\},$$

which becomes more intuitive if one states it in diagrammatic form: a morphism $f \rightarrow g$ in $\mathcal{C}^\rightarrow$ is defined to be a commutative square in $\mathcal{C}$ of the form

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

Given such a morphism $f \rightarrow g$, we can compose it with another morphism $g \rightarrow h$ by composing the vertical arrows in the two squares. Since associativity and existence of identities are clear, we see that $\mathcal{C}^\rightarrow$ is indeed a category. We already saw the arrow category $\mathbf{Vect}_K^\rightarrow$ in Example 2.6.7.

Example 2.8.6. Continuing on from Example 2.8.4, what does a morphism of diagrams amount to in that case?

The answer is that we get the notion of morphism of pointed successor sets which implicitly appears in Theorem 1.1.2, namely that it is a diagram of the form

$$\begin{array}{ccc} 1 & \xrightarrow{p} & A \\ \text{id}_1 \downarrow & & \downarrow h \\ 1 & \xrightarrow{q} & B \end{array} \quad \begin{array}{c} \curvearrowright^t \\ \curvearrowright^u \end{array}$$

where partial commutativity in the form $q = hp$ and $uh = ht$ is required. The fact that id_1 appears on the left is clear, because the identity is the only map from a singleton set to itself.

The equation $q = hp$ amounts to saying that h should respect the initial state, while $uh = ht$ means that h should respect the dynamics. This is closely related to core concepts of physics and computer science.¹¹

2.9 Functor categories

As we saw Section 2.7, we can consider a functor $F : \mathcal{C} \rightarrow \mathcal{D}$ as a morphism in $\mathbf{Cat}$, the category of small categories. But we can also consider it as an *object* in the **functor category** $[\mathcal{C}, \mathcal{D}]$, where the objects are functors $\mathcal{C} \rightarrow \mathcal{D}$ and the morphisms $F \rightarrow G$ are the natural transformations $F \Rightarrow G$, with composition being vertical composition. You may want to convince yourself that vertical composition is indeed associative because of associativity of composition in $\mathcal{D}$, and that the identity natural transformation $\text{id}_F : F \Rightarrow F$, all of whose components are identity morphisms, acts as the identity morphism of $F \in [\mathcal{C}, \mathcal{D}]$.

Example 2.9.1 (Arrow category, II). Let $\mathcal{C}$ be the category consisting of two objects and only one non-identity morphism:

$$\bullet_0 \xrightarrow{\text{wa}} \bullet_1$$

Since the only possible compositions are those with the identity morphisms of the two objects, it is clear how to compose morphisms and that this $\mathcal{C}$ is indeed a category. We write wa for the non-identity morphism since it is sometimes called “walking arrow”, in the sense of the purest instantiation of an arrow (morphism) in a category.¹²

¹¹In physics, we can for example define a *conserved quantity* as a morphism $h : A \rightarrow B$ where $u = \text{id}_B$, meaning that the quantity does not depend on time. In computer science, see e.g. the concept of simulation between automata or state transition systems, as defined at [en.wikipedia.org/wiki/Simulation_\(computer_science\)](https://en.wikipedia.org/wiki/Simulation_(computer_science)).

¹²Compare with standard English phrases like “walking encyclopedia” for a person who knows a lot, or “walking disaster” for a person who is very clumsy or unlucky.

Saying that $\mathbf{wa}$ is the “purest instantiation” of a morphism in a category is justified by the fact that for any category $\mathcal{D}$, the functors $F : \mathcal{C} \rightarrow \mathcal{D}$ are in natural correspondence with morphisms in $\mathcal{D}$: for any morphism $g : A \rightarrow X$ in $\mathcal{D}$, we obtain a functor $F_g : \mathcal{C} \rightarrow \mathcal{D}$ by mapping the two objects of $\mathcal{C}$ to A and X , respectively, and $\mathbf{wa}$ to g . It is clear that every functor $F : \mathcal{C} \rightarrow \mathcal{D}$ arises in this way for a unique morphism $g : X \rightarrow Y$. In other words, just as in Example 2.8.5, the functors $\mathcal{C} \rightarrow \mathcal{D}$ “are” the morphisms of $\mathcal{D}$.

Now if we have two morphisms $g : A \rightarrow X$ and $h : B \rightarrow Y$ in $\mathcal{D}$, what are the natural transformations $F_g \Rightarrow F_h$? Such a transformation has two components given by the vertical arrows in the diagram

$$\begin{array}{ccc} A & \xrightarrow{g} & X \\ \alpha_0 \downarrow & & \downarrow \alpha_1 \\ B & \xrightarrow{h} & Y \end{array} \quad (2.18)$$

which is the naturality square of the natural transformation at the morphism $\mathbf{wa}$. Since the naturality squares for the two identity morphisms commute trivially, the commutation of this square in the form $\alpha_1 g = h \alpha_0$ is all that is needed. So to summarize, a morphism $g \rightarrow h$ in the functor category $[\mathcal{C}, \mathcal{D}]$ is a commutative square (2.17) in $\mathcal{D}$.

The composition of two such morphisms is given by “composing the squares”,

$$\begin{array}{ccc} A & \xrightarrow{g} & X \\ \alpha_0 \downarrow & & \downarrow \alpha_1 \\ B & \xrightarrow{h} & Y \\ \beta_0 \downarrow & & \downarrow \beta_1 \\ C & \xrightarrow{k} & Z \end{array} \quad (2.19)$$

since this precisely implements the vertical composition of the natural transformations. While we do already know that the composite square commutes because the vertical composition of natural transformations is a natural transformation again, it is instructive to verify this again by a simple diagram chase,

$$\beta_1 \alpha_1 g = \beta_1 h \alpha_0 = k \beta_0 \alpha_0.$$

Example 2.9.2. Let us consider $\mathcal{C} = \mathbf{BN}$, the category with a single object $*$ having $(\mathbb{N}, +)$ as its endomorphism monoid. For reasons that will become clear, we write $\mathbf{we}^{(n)}$ for these morphisms, so that $\mathbf{we}^{(0)}$ is the identity morphism and composition is given by

$$\mathbf{we}^{(m)} \circ \mathbf{we}^{(n)} = \mathbf{we}^{(m+n)}.$$

Note that $\mathbf{we}^{(n)}$ is the n -fold composition of $\mathbf{we} := \mathbf{we}^{(1)}$ with itself.

Then what is a functor $F : \mathbf{BN} \rightarrow \mathcal{D}$? It consists of a choice of object $F(*)$ in $\mathcal{D}$ and a morphism $F(\mathbf{we}) : F(*) \rightarrow F(*)$. Then by functoriality, F must map $\mathbf{we}^{(n)}$ to the n -fold composition of $F(\mathbf{we})$ with itself. In other words, we can identify the functors $\mathbf{BN} \rightarrow \mathcal{D}$ with pairs (X, f) consisting of an object X in $\mathcal{D}$ and an endomorphism $f : X \rightarrow X$. Based on this identification, we also think of $\mathbf{we}$ as the “walking endomorphism”.

Such pairs (X, f) can also be thought of as a **dynamical system** in $\mathcal{D}$, meaning an object X together with a morphism $f : X \rightarrow X$ which specifies how it evolves in one step of time, but now without a specified initial state (in contrast to Example 2.8.4). Applying the corresponding functor $\mathbf{BN} \rightarrow \mathcal{D}$ to the morphism $\mathbf{we}^{(n)}$ then gives the time evolution of the system in n steps.

Now what is a natural transformation between two such functors? Since $\mathbf{BN}$ has only one object, a natural transformation between two such functors is simply a morphism in $\mathcal{D}$. Indeed if we are given dynamical systems (X, f) and (Y, g) , then a natural transformation between the

corresponding functors is a morphism $h : X \rightarrow Y$ in $\mathcal{D}$ such that the diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & X \\ h \downarrow & & \downarrow h \\ Y & \xrightarrow{g} & Y \end{array}$$

commutes, which is precisely the naturality square for the morphism $\mathbf{we}$. The other naturality square now commute automatically, as for example for $\mathbf{we}^{(2)} = \mathbf{we} \circ \mathbf{we}$, it corresponds to the outer square in

$$\begin{array}{ccccc} X & \xrightarrow{f} & X & \xrightarrow{f} & X \\ h \downarrow & & \downarrow h & & \downarrow h \\ Y & \xrightarrow{g} & Y & \xrightarrow{g} & Y \end{array}$$

As with (2.18), the commutativity of the outer square follows as a consequence of commutativity of each of the smaller squares.

What we have now shown can also be summarized by saying that the functor category $[\mathbf{BN}, \mathcal{C}]$ is isomorphic to the category of diagrams in $\mathcal{C}$ whose shape is the graph with a single loop:



3 Universal properties

In category theory, we study an object only through how it interacts with other objects. In order to show that this really is a meaningful endeavour, we need to convince ourselves that this is powerful enough to *characterize* objects. This is achieved elegantly by the concept of universal property. We approach this concept by talking about representable functors first.

3.1 Representable functors

Definition 3.1.1. A functor $F : \mathcal{C} \rightarrow \mathbf{Set}$ is **representable** if there is an object A and a natural isomorphism

$$F \cong \mathcal{C}(A, -). \quad (3.1)$$

Here, $\mathcal{C}(A, -) : \mathcal{C} \rightarrow \mathbf{Set}$ is the hom-functor of A in the sense of Example 2.6.6. Representability is often used in order to *define* the object A in terms of the functor F . In particular, we will see that the representability characterizes A essentially uniquely. But before thinking about this, let us consider some examples on which we can gain some intuition.

Example 3.1.2. The forgetful functor $U : \mathbf{Grp} \rightarrow \mathbf{Set}$ from groups to sets is representable.

To see this, note that for every group G , the map

$$\begin{aligned} \mathbf{Grp}(\mathbb{Z}, G) &\longrightarrow G \\ \varphi &\longmapsto \varphi(1) \end{aligned} \quad (3.2)$$

is a bijection, since every group homomorphism $\varphi : \mathbb{Z} \rightarrow G$ is uniquely determined by $\varphi(1)$, which shows that the map is injective. For surjectivity, note that every element $g \in G$ gives rise to a group homomorphism $\mathbb{Z} \rightarrow G$ by sending $n \mapsto g^n$.¹³

Furthermore, (3.2) is natural in G , meaning that for every group homomorphism $f : G \rightarrow H$, the following diagram commutes:

$$\begin{array}{ccc} \mathbf{Grp}(\mathbb{Z}, G) & \xrightarrow{\varphi \mapsto \varphi(1)} & G \\ f \circ - \downarrow & & \downarrow f \\ \mathbf{Grp}(\mathbb{Z}, H) & \xrightarrow{\psi \mapsto \psi(1)} & H \end{array}$$

On the left, we use that the hom-functor $\mathbf{Grp}(\mathbb{Z}, -)$ acts on morphisms by composition. The diagram commutes essentially trivially because of $(f \circ \varphi)(1) = f(\varphi(1))$.

Therefore (3.2) is a natural isomorphism which witnesses the representability of U .

Example 3.1.3. For field K , let us consider the **tensor product** of vector spaces $U, V \in \mathbf{Vect}_K$. By definition, this is a vector space $U \otimes V$ together with a bilinear map

$$\begin{aligned} U \times V &\longrightarrow U \otimes V \\ (u, v) &\longmapsto u \otimes v \end{aligned} \quad (3.3)$$

which is universal in the sense that for every vector space W and bilinear map $b : U \times V \rightarrow W$, there is a unique linear map $\hat{b} : U \otimes V \rightarrow W$ such that the diagram

$$\begin{array}{ccc} U \times V & \xrightarrow{\quad \otimes \quad} & U \otimes V \\ & \searrow b \quad \swarrow \hat{b} & \\ & W & \end{array}$$

commutes, or in other words such that $\hat{b}(u \otimes v) = b(u, v)$ for all $u \in U$ and $v \in V$.

¹³This can also be phrased as the statement that $\mathbb{Z}$ is a *free group* on one generator.

What does this have to do with representability? To see this, put

$$\text{Bil}_K(U, V; W) := \{\text{bilinear maps } U \times V \rightarrow W\}.$$

Now if $f : W \rightarrow W'$ is linear, post-composing any bilinear map $U \times V \rightarrow W$ with f produces a bilinear map $U \times V \rightarrow W'$. Based on this, it is easy to see that we are dealing with a functor

$$\text{Bil}_K(U, V; -) : \text{Vect}_K \longrightarrow \text{Set}.$$

This functor is representable, since by definition we have a bijection

$$\text{Vect}_K(U \otimes V, W) \xrightarrow{\cong} \text{Bil}_K(U, V; W)$$

given by pre-composing with the bilinear map (3.3). Make sure to verify that this is natural in W !

Based on this, one views (3.3) as the “universal” bilinear map out of U and V : every other bilinear map factors through it in a unique way! This is an example of a universal property. Our goal will now be to understand how the concept of representability encodes properties like this in general.

3.2 The Yoneda lemma

The Yoneda lemma is going to be our first nontrivial “theorem” in category theory. As with many of the basic results of category theory, it is a statement that is abstract and may take some time to understand, but at the same time quite straightforward to prove.

Lemma 3.2.1 (Yoneda). *Given a locally small category $\mathcal{C}$, an object $A \in \mathcal{C}$, and a functor $F : \mathcal{C} \rightarrow \text{Set}$, there is a bijection*

$$\begin{aligned} [\mathcal{C}, \text{Set}](\mathcal{C}(A, -), F) &\xrightarrow{\cong} F(A) \\ \alpha &\longmapsto \alpha_A(\text{id}_A). \end{aligned} \tag{3.4}$$

Keep in mind that the brackets $[\mathcal{C}, \text{Set}]$ denote the functor category. Thus the statement establishes a bijection between the natural transformations $\mathcal{C}(A, -) \Rightarrow F$ and the elements of $F(A)$. The assumption of $\mathcal{C}$ being locally small is relevant only for ensuring that the hom-functor $\mathcal{C}(A, -)$ is a functor into Set .

In addition to the importance of the statement itself, the formula (3.5) that we develop in the proof will also play a role for our understanding of universal properties.

Proof. We first show the following auxiliary statement: for every $\alpha : \mathcal{C}(A, -) \Rightarrow F$ and every $f : A \rightarrow X$, we have

$$\boxed{\alpha_X(f) = F(f)(\alpha_A(\text{id}_A))}. \tag{3.5}$$

This formula is best thought of together with the naturality square

$$\begin{array}{ccc} \mathcal{C}(A, A) & \xrightarrow{f \circ -} & \mathcal{C}(A, X) \\ \alpha_A \downarrow & & \downarrow \alpha_X \\ F(A) & \xrightarrow{F(f)} & F(X) \end{array}$$

which also provides the proof: starting with id_A in the top left, we get $\alpha_X(f)$ in the bottom right by taking the upper path, and $F(f)(\alpha_A(\text{id}_A))$ by taking the lower path, so that (3.5) follows by commutativity.

We now turn to the actual claim about (3.4). Concerning injectivity, (3.5) already tells us that α is uniquely determined by $\alpha_A(\text{id}_A)$, which proves injectivity. More formally, assume that

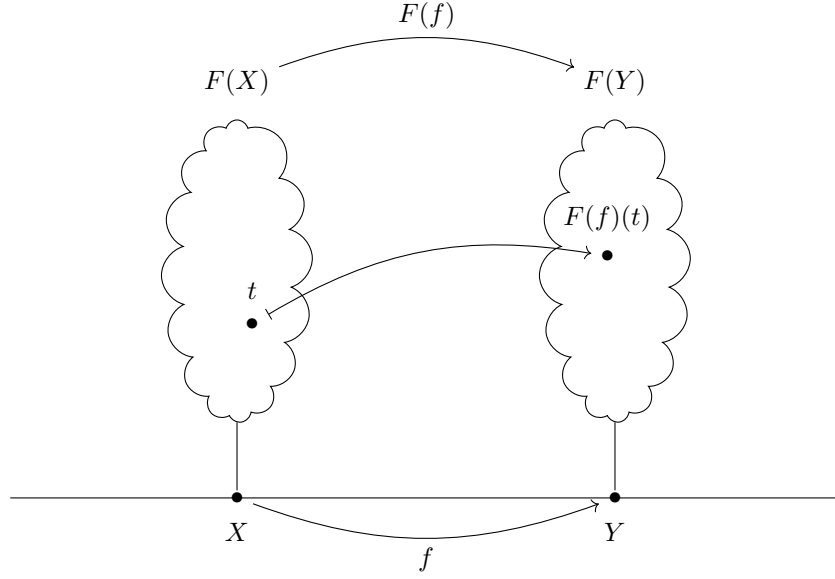


Figure 2: How to visualize a functor $F : \mathcal{C} \rightarrow \mathbf{Set}$. The “ground” is the category $\mathcal{C}$ with its objects and morphisms, and above every object $X \in \mathcal{C}$ we have a “cloud” or “canopy” representing the set $F(X)$. Every morphism $f : X \rightarrow Y$ lifts to a map $F(f) : F(X) \rightarrow F(Y)$ between the canopies.

$\alpha, \beta : \mathcal{C}(A, -) \Rightarrow F$ are two natural transformations such that $\alpha_A(\text{id}_A) = \beta_A(\text{id}_A)$. Then by (3.5), we have

$$\alpha_X(f) = F(f)(\alpha_A(\text{id}_A)) = F(f)(\beta_A(\text{id}_A)) = \beta_X(f),$$

as was to be shown.

For surjectivity, take any $s \in F(A)$. Then (3.5) suggests how we should find an α such that $\alpha_A(\text{id}_A) = s$, namely by defining

$$\alpha_X(f) := F(f)(s)$$

for every $f \in \mathcal{C}(A, X)$ for every X . If we can prove that this α is a natural transformation, then we are done, because we have $\alpha_A(\text{id}_A) = s$ by construction. Hence we need to check that for every $g : X \rightarrow Z$, the naturality square

$$\begin{array}{ccc} \mathcal{C}(A, X) & \xrightarrow{g \circ -} & \mathcal{C}(A, Z) \\ \alpha_X \downarrow & & \downarrow \alpha_Z \\ F(X) & \xrightarrow{F(g)} & F(Z) \end{array}$$

commutes. Starting with any $f \in \mathcal{C}(A, X)$ in the top left, this follows by

$$\begin{aligned} \alpha_Z(gf) &= F(gf)(s) \\ &= F(g)(F(f)(s)) \\ &= F(g)(\alpha_X(f)). \end{aligned}$$

as was to be shown. □

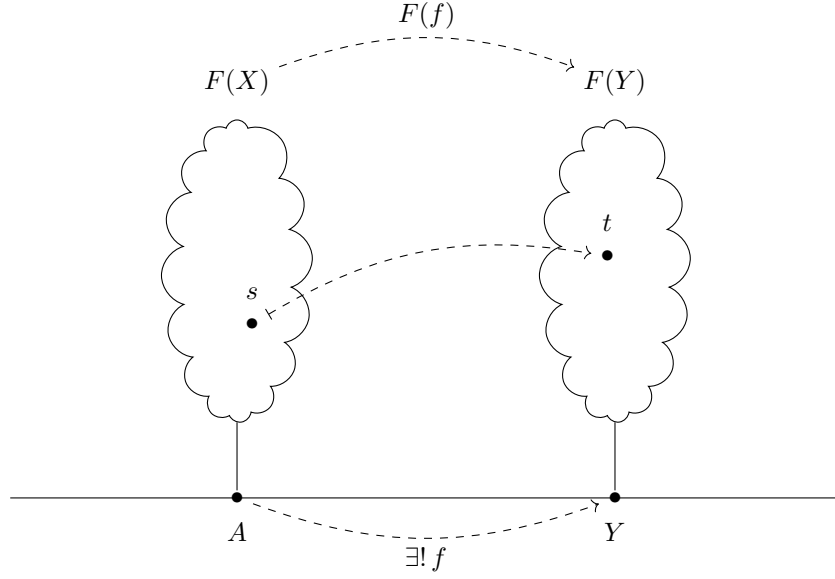


Figure 3: Visualization of a universal property: A is the universal object, s is the universal element, and the universal property states that for every X and every $t \in F(X)$, there is a unique morphism $f : A \rightarrow X$ such that $t = F(f)(s)$.

3.3 Universal elements

In Example 3.1.3, we had derived representability from the unique factorization property of bilinear maps across the tensor product. This is how it works in general! Before turning to the general result, consider the visual way to think about a functor $F : \mathcal{C} \rightarrow \mathbf{Set}$ that is offered by Figure 2.

Proposition 3.3.1. *A functor $F : \mathcal{C} \rightarrow \mathbf{Set}$ is representable if and only if there is an object $A \in \mathcal{C}$ and an element $s \in F(A)$ such that for every $X \in \mathcal{C}$ and every $t \in F(X)$, there is a unique $f : A \rightarrow X$ such that*

$$F(f)(s) = t.$$

We call such an $s \in F(A)$ a **universal element** of F , and the unique existence of f a **universal property** of the object A . See Figure 3 for a visualization of this universal property.

Proof. By the Yoneda lemma, we already know that the natural transformations $\mathcal{C}(A, -) \Rightarrow F$ correspond to the elements of $F(A)$ via the formula (3.5). So it remains to be shown that α is a natural *isomorphism* if and only if the stated condition holds. But this holds precisely if every α_X is bijective. In light of (3.5), this is exactly the condition that we are now assuming. $\square$

Remark 3.3.2. An inspection of the proof via (3.5) shows that the universal element $s \in F(A)$ is precisely the counterpart of the identity morphism $\text{id}_A \in \mathcal{C}(A, A)$ under the natural isomorphism $\mathcal{C}(A, -) \Rightarrow F$.

Example 3.3.3. For the functor $\text{Bil}_{\mathbb{K}}(U, V; -)$ from Example 3.1.3, the universal element is precisely the bilinear map

$$\begin{aligned} U \times V &\longrightarrow U \otimes V \\ (u, v) &\longmapsto u \otimes v. \end{aligned}$$

Indeed the universal property states that for every vector space W and bilinear map $b : U \times V \rightarrow W$, there is a unique linear map $\hat{b} : U \otimes V \rightarrow W$ such that the diagram

$$\begin{array}{ccc} U \times V & \xrightarrow{\quad \otimes \quad} & U \otimes V \\ & \searrow b \quad \swarrow \hat{b} & \\ & W & \end{array}$$

commutes. But this is precisely the “unique factorization” we had already seen in Example 3.1.3.

Example 3.3.4. In the case of Example 3.1.2, a universal element is $1 \in \mathbb{Z}$, and the universal property is that for every group G and every $g \in G$, there is a unique group homomorphism $\varphi : \mathbb{Z} \rightarrow G$ such that $\varphi(1) = g$.

Example 3.3.5 (Quotient topology). Let T be a topological space and $\sim \subseteq T \times T$ an equivalence relation on T . Then the **quotient space** $T/\sim$ can be constructed as the set of equivalence classes $[t]$ of $t \in T$, where a subset $U \subseteq T/\sim$ is open if and only if its preimage $\pi^{-1}(U)$ under the canonical projection

$$\begin{aligned} \pi : T &\longrightarrow T/\sim \\ t &\longmapsto [t] \end{aligned}$$

is open in T . In this example, we describe the universal property enjoyed by this quotient space.

Let $F : \mathbf{Top} \rightarrow \mathbf{Set}$ be the functor that acts on objects as

$$F(X) := \{f : T \rightarrow X \text{ continuous} \mid t \sim t' \Rightarrow f(t) = f(t')\},$$

with the action on morphisms given by composition: for any any continuous $f : T \rightarrow X$ that preserves the equivalence as stated and any continuous $g : X \rightarrow Y$, also $g \circ f$ preserves the equivalence, and therefore

$$\begin{aligned} F(g) : F(X) &\longrightarrow F(Y) \\ f &\longmapsto g \circ f \end{aligned}$$

is well-defined.

Let us now use Proposition 3.3.1 to argue that F is representable. To this end, we show that the projection $\pi : T \rightarrow T/\sim$, which belongs to $F(T/\sim)$, is a universal element. The universal property states that if $f : T \rightarrow X$ is any continuous map which preserves the equivalence relation, then there is a unique continuous map $\hat{f} : T/\sim \rightarrow X$ such that the diagram

$$\begin{array}{ccc} T & \xrightarrow{\quad \pi \quad} & T/\sim \\ & \searrow f \quad \swarrow \hat{f} & \\ & X & \end{array}$$

commutes. You should convince yourself that the definition of the quotient topology on $T/\sim$ indeed makes this statement true!

In general, to what extent are the universal object A and the universal element s unique? Here’s the answer:

Proposition 3.3.6. *Let $F : \mathcal{C} \rightarrow \mathbf{Set}$ be a representable functor with universal elements $s \in F(A)$ and $t \in F(B)$. Then there is a unique isomorphism $f : A \rightarrow B$ such that $F(f)(s) = t$.*

This is why one says that the object A is **unique up to unique isomorphism**. And this is also what we meant by “essentially unique” after Definition 3.1.1.

Proof. By the universal property of s , we know that there is a unique morphism $f : A \rightarrow B$ such that $F(f)(s) = t$. Similarly by the universal property of t , we know that there is a unique morphism $g : B \rightarrow A$ such that $F(g)(t) = s$. Then we have by functoriality

$$F(gf)(s) = F(g)(F(f)(s)) = F(g)(t) = s.$$

But then again by the universal property of s , we get that $gf = \text{id}_A$. Similarly, we can show that $fg = \text{id}_B$. $\square$

Example 3.3.7. Let us continue with the forgetful functor $U : \mathbf{Grp} \rightarrow \mathbf{Set}$ from Example 3.1.2. We already noted that $1 \in \mathbb{Z}$ is a universal element, and that the universal property states that for every group G and every $g \in G$, there is a unique group homomorphism $\varphi : \mathbb{Z} \rightarrow G$ such that $\varphi(1) = g$. One also says that this makes $\mathbb{Z}$ into the *free group on one generator*, where the generator is the universal element 1. But we can also take $-1 \in \mathbb{Z}$ as a universal element: for every group G and every $g \in G$, there is *also* a unique group homomorphism $\psi : \mathbb{Z} \rightarrow G$ such that $\psi(-1) = g$, namely the one given by $\psi(n) = g^{-n}$. This is consistent with Proposition 3.3.6, where the isomorphism $f : \mathbb{Z} \rightarrow \mathbb{Z}$ is given by $f(n) = -n$. We think about this as saying that 1 and -1 are really the same universal element, just with different labels, as they only differ by the “relabelling” isomorphism f .

No other element of $\mathbb{Z}$ is universal: for example $2 \in \mathbb{Z}$ is not universal, because for any homomorphism $\eta : \mathbb{Z} \rightarrow G$, the element $\eta(2) \in G$ cannot be arbitrary, but it has the special property that there exists a square root in G . Since there do exist groups in which not every element has a square root (for example just $1 \in \mathbb{Z}$ again), this shows that 2 is not universal.

Example 3.3.8. In the case of the quotient topology from Example 3.3.5, we already saw that the quotient map $\pi : T \rightarrow T/\sim$ is a universal element. At this level of generality we cannot describe other universal elements explicitly, but this arguably also is not necessary, because the universal property itself is all that we need to know in order to work with the quotient space. In more detail, a topological space Q together with any continuous map $p : T \rightarrow Q$ displays the universal property if the following hold:

- ▷ p respects the equivalence relation, and
- ▷ if $f : T \rightarrow X$ is any continuous map which respects the equivalence relation, then there is a unique continuous map $\hat{f} : Q \rightarrow X$ such that the diagram

$$\begin{array}{ccc} T & \xrightarrow{p} & Q \\ & \searrow f & \swarrow \hat{f} \\ & X & \end{array}$$

commutes.

These properties are everything that we use in mathematical practice about the quotient topology! For example, we can use it to show that a set $U \subseteq Q$ is open if and only if $p^{-1}(U)$ is open in T .¹⁴ What we do not use—and what is actually completely irrelevant for mathematical practice—is whether the points of Q are equivalence classes of points of T or something else. This is analogous to how we had noted around (1.1) that it is completely irrelevant for mathematical practice what the natural numbers “are”. For both the quotient topology and the natural numbers, it only matters how they *behave*, or more precisely how they *interact* with other objects, and this is what the universal property captures.

A very particular case of a functor which may or may not be representable is $F : \mathcal{C} \rightarrow \mathbf{Set}$ given by $F(X) := 1$ for every $X \in \mathcal{C}$, where 1 is any one-element set. The action of F on morphisms

¹⁴See the homework!

is automatically clear: F must assign to every morphism in $\mathcal{C}$ the unique map $1 \rightarrow 1$, and the functoriality is obvious because of the uniqueness. When is such a functor representable?

To answer this, let us spell out what the universal property amounts of Proposition 3.3.1 amounts to in this case. Since the elements s and t always exist uniquely, what we get is simply the following: the universal object A is such that for every $X \in \mathcal{C}$, there is a unique morphism $f : A \rightarrow X$. It is useful to have a name for this property.

Definition 3.3.9. An object A in a category $\mathcal{C}$ is called **initial** if for every object $X \in \mathcal{C}$, there is a unique morphism $f : A \rightarrow X$.

Example 3.3.10. In **Set** and **Top**, the empty set $\emptyset$ (or empty space) is initial, because for every X there is exactly one map $\emptyset \rightarrow X$, namely the empty one, and in the case of **Top** it is trivially continuous.

Example 3.3.11 (Universal property of $\mathbb{N}$). With the definition of pointed successor set from Definition 1.1.1 in mind, let us consider pointed successor sets (A, p, t) as a category, where a morphism $(A, p, t) \rightarrow (B, q, u)$ is defined as a map $f : A \rightarrow B$ which makes the diagram

$$\begin{array}{ccccc} & & A & \xrightarrow{t} & A \\ & p \nearrow & \downarrow f & & \downarrow f \\ 1 & & B & \xrightarrow{u} & B \\ & q \searrow & & & \end{array}$$

commute. These morphisms compose in the obvious manner, and thereby we obtain a category **PSuccSet**. The recursion principle of Theorem 1.1.2 now states precisely that $(\mathbb{N}, 0, s)$ is an initial object in **PSuccSet**. This is the universal property of $\mathbb{N}$ as a pointed successor set, and it is how we can *define* $\mathbb{N}$ in the first place. This definition encapsulates the idea of recursive constructions (including inductive proofs). Just like before, it is irrelevant what the elements of $\mathbb{N}$ are; what matters for mathematical practice is the universal property.

The universal properties that we have considered so far are all “**mapping out**” **universal properties**, because the representability tells us how to map out of the universal object, and we saw that this characterizes the universal object essentially uniquely by the Yoneda lemma. If we apply these same concepts to opposite categories, then we get “**mapping in**” **universal properties**, which characterize a universal object by saying how to map into it.

For example, a **terminal object** in a category $\mathcal{C}$ is an object A such that for every object $X \in \mathcal{C}$, there is a unique morphism $f : X \rightarrow A$. Such an object represents the functor $\mathcal{C}^{\text{op}} \rightarrow \mathbf{Set}$ which maps everything to 1, and equivalently is an initial object in the opposite category $\mathcal{C}^{\text{op}}$.

Example 3.3.12. (a) Any singleton set is a terminal object in **Set**.

(b) The *zero ring* Z is any ring in which the two neutral elements coincide, meaning $0 = 1$. Such a ring contains only the zero element, because every element x satisfies

$$x = 1 \cdot x = 0 \cdot x = 0.$$

On the other hand, it is clear that $Z = \{0\}$ indeed satisfies the axioms of a commutative ring, because all axioms simply reduce to $0 = 0$.

Now Z is terminal in **CRng**, because for any commutative ring R , the unique map $R \rightarrow Z$ is trivially a homomorphism, since again both the additivity and the multiplicativity axioms reduce to $0 = 0$.

(c) Similarly, any group consisting only of the neutral element, that is $G = \{e\}$, is a terminal object in **Grp**.

Example 3.3.13 (Cartesian products). Given sets A and B , their **Cartesian product** is usually defined as the set of pairs,

$$A \times B := \{(a, b) \mid a \in A, b \in B\}.$$

As we will see now, this set can also be characterized by a universal property in **Set**. Once again, this property captures exactly how one uses the Cartesian product in practice. In Section 4, we will learn that this is a special case of the much more general concept of limit.

So how do we specify a map from a third set X into the Cartesian product $A \times B$? The answer is that we need to map *to each component*: we specify maps $f : X \rightarrow A$ and $g : X \rightarrow B$, and then we can *pair* these maps to obtain a map

$$\begin{aligned} (f, g) : X &\longrightarrow A \times B \\ x &\longmapsto (f(x), g(x)). \end{aligned}$$

To formulate this as a universal property, we observe that (f, g) is the *unique* map $X \rightarrow A \times B$ whose “components” are f and g , and we still need to make precise what this means. To this end, we consider the functor

$$\begin{aligned} F : \mathbf{Set}^{\mathrm{op}} &\longrightarrow \mathbf{Set} \\ X &\longmapsto \mathbf{Set}(X, A) \times \mathbf{Set}(X, B), \end{aligned}$$

where the action on morphisms is given by composition in each component. And indeed this is what we need: if we write $A \times B$ for a universal object, then a universal element is given by a pair (p_A, p_B) of **product projections**

$$p_A : A \times B \longrightarrow A, \quad p_B : A \times B \longrightarrow B,$$

and the universal property states that for every two maps $f : X \rightarrow A$ and $g : X \rightarrow B$, there is a unique map $(f, g) : X \rightarrow A \times B$ such that

$$f = p_A \circ (f, g) \quad \text{and} \quad g = p_B \circ (f, g).$$

This implements the idea that (f, g) is uniquely determined by the requirement that its two components should be f and g , where we now understand that the “components” of a map $X \rightarrow A \times B$ are the two maps $X \rightarrow A$ and $X \rightarrow B$ that we obtain by composing with the product projections p_A and p_B .

Let us consider the universal property with $X = 1$ a singleton set. Then the statement is that every pair of elements written in the form $a : 1 \rightarrow A$ and $b : 1 \rightarrow B$ gives rise to an element $(a, b) : 1 \rightarrow A \times B$, which is the unique element whose two components are a and b . Like this, we even recover the fact that the elements of the product $A \times B$ are pairs (a, b) !¹⁵

As a final comment, let us note that the universal property of the Cartesian product is secretly already known to everyone from school: in the special case $A = B = \mathbb{R}$, and where $X = 1$ is a singleton set, it specializes to the statement that the elements of $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ are uniquely determined by their two components, which are the two coordinates. In terms of a more informal slogan, one can perhaps formulate this insight as¹⁶

Cartesian coordinates = universal property of the Cartesian product.

Let us end this subsection with two further examples of universal properties which come in *both* a “mapping in” variant and a “mapping out” variant.

¹⁵Though note that it plays no role whether these elements actually *are* pairs in any formal sense. So more accurately phrased, the universal property lets us *label* the elements of $A \times B$ by pairs (a, b) .

¹⁶Although we have considered the case of $\mathbb{R}^2$, the generalization to $\mathbb{R}^n$ is straightforward, and you may want to work it out for yourself.

Example 3.3.14 (Kernels and cokernels of linear maps). We return to the formation of kernels of linear maps. For a linear map $f : V \rightarrow W$ between vector spaces, the kernel $\ker(f)$ is a subspace of V . Can it be characterized by a universal property?

The answer is indeed yes, namely the following: for any $U \in \mathbf{Vect}_K$ and any linear map $g : U \rightarrow V$ with $fg = 0$, there is a unique linear map $\hat{g} : U \rightarrow \ker(f)$ such that the diagram

$$\begin{array}{ccccc} & & U & & \\ & \swarrow \exists! \hat{g} & & \searrow g & \\ \ker f & \hookrightarrow & V & \xrightarrow{f} & W \end{array}$$

commutes, where the arrow $\ker(f) \hookrightarrow V$ is the inclusion map. In other words, this universal property says that the inclusion $\ker(f) \hookrightarrow V$ is the universal element of the functor

$$\begin{aligned} F : \mathbf{Vect}_K^{\text{op}} &\longrightarrow \mathbf{Set} \\ U &\longmapsto \{\text{linear maps } g : U \rightarrow V \text{ with } fg = 0\}, \end{aligned}$$

with the action on morphisms given by composition as usual. Once again, this universal property is secretly how we work with the kernel in practice. Applying it to morphisms of the form $g : K \rightarrow V$, which we already know to correspond to the elements of V , shows that the elements of $\ker(f)$ are in bijection with the elements $x \in V$ with $f(x) = 0$, which is the element-based definition of the kernel that we are used to.

Okay, what happens if we dualize this universal property to a “mapping out” property? That is, what about looking at maps g out of the codomain of f rather than into its domain? This amounts to considering the functor

$$\begin{aligned} F : \mathbf{Vect}_K &\longrightarrow \mathbf{Set} \\ U &\longmapsto \{\text{linear maps } g : W \rightarrow U \text{ with } gf = 0\}. \end{aligned}$$

Assuming that it exists, let us write $\text{coker}(f)$ for the corresponding universal object and call it the **cokernel** of f . A universal element is now a linear map $W \twoheadrightarrow \text{coker}(f)$ such that if $g : W \rightarrow U$ is any linear map satisfying $gf = 0$, then there is a unique linear map $\hat{g} : \text{coker}(f) \rightarrow U$ such that the diagram

$$\begin{array}{ccccc} V & \xrightarrow{f} & W & \twoheadrightarrow & \text{coker}(f) \\ & & \searrow g & & \swarrow \exists! \hat{g} \\ & & U & & \end{array} \tag{3.6}$$

commutes. This universal property is satisfied if we use the quotient space

$$\text{coker}(f) := W/\text{im}(f),$$

and take the universal map $W \twoheadrightarrow W/\text{im}(f)$ to be the projection. To see this, note that defining

$$\hat{g}([w]) := g(w)$$

makes $\hat{g}$ well-defined thanks to the assumption that $gf = 0$,¹⁷ and the triangle in (3.6) then commutes by definition. The uniqueness is clear because the projection $W \twoheadrightarrow W/\text{im}(f)$ is surjective and hence an epimorphism, which gives us the uniqueness.

¹⁷Work this out formally if it's not completely clear to you! If $w, w' \in W$ are representatives of the same equivalence class $[w] = [w']$, then why is $g(w) = g(w')$?

4 Limits and colimits

4.1 Categorical products

We have seen in Example 3.3.13 that the Cartesian product of two sets enjoys a meaningful universal property which characterizes it. It is quite straightforward to generalize this to *any* category, where we therefore obtain a notion of taking the “product” of two objects.

Definition 4.1.1. *Given a category $\mathcal{C}$ and objects $A, B \in \mathcal{C}$, their **categorical product** is an object $A \times B \in \mathcal{C}$ together with morphisms*

$$p_A : A \times B \longrightarrow A, \quad p_B : A \times B \longrightarrow B$$

such that for every object $X \in \mathcal{C}$ and every pair of morphisms $f : X \rightarrow A$ and $g : X \rightarrow B$, there is a unique morphism $(f, g) : X \rightarrow A \times B$ such that the diagram

$$\begin{array}{ccccc}
 & & f & & \\
 & & \curvearrowright & & \\
 & & & & A \\
 X & \overset{(f,g)}{\dashrightarrow} & A \times B & \xrightarrow{p_A} & \\
 & & \curvearrowleft & & \\
 & & g & & B
 \end{array}$$

*commutes. We call the morphisms p_A and p_B the **product projections** and the morphism (f, g) the **pairing** of f and g .*

Remark 4.1.2. Let us see how this definition indeed amounts to a universal property in the sense of the previous section. To this end, we temporarily write $[f, g]$ for the pairing of f and g , in order not to confuse the pairing with the set-theoretic pair (f, g) .

We then argue that the categorical product $A \times B$ is a representing object for the functor

$$\begin{aligned}
 \mathcal{C}^{\text{op}} &\longrightarrow \text{Set} \\
 X &\longmapsto \mathcal{C}(X, A) \times \mathcal{C}(X, B) \\
 (f : X \rightarrow Y) &\longmapsto \left(\begin{array}{c} \mathcal{C}(Y, A) \times \mathcal{C}(Y, B) \longrightarrow \mathcal{C}(X, A) \times \mathcal{C}(X, B) \\ (g, h) \longmapsto (gf, hf) \end{array} \right).
 \end{aligned} \tag{4.1}$$

This matches up with Definition 4.1.1 via Proposition 3.3.1 as follows. The pair (p_A, p_B) is the universal element of the functor at the universal object $A \times B$. The condition that an arbitrary element (f, g) must be obtained by pushing forward the universal element (p_A, p_B) along a unique morphism $[f, g] : X \rightarrow A \times B$ amounts to the equations

$$f = p_A \circ [f, g] \quad \text{and} \quad g = p_B \circ [f, g],$$

and these express precisely the commutativity of the diagram.

The essential uniqueness of the universal object from Proposition 3.3.6 justifies speaking of *the* categorical product.

Remark 4.1.3. There is no circularity with respect to the appearance of the symbol “ $\times$ ” used in order to define “ $\times$ ”. This is because the “ $\times$ ” in (4.1) is the Cartesian product of sets, while the “ $\times$ ” in $A \times B$ denotes the categorical product of objects in $\mathcal{C}$. We thus use the former in order to define the latter.

Example 4.1.4. Let $A, B \in \mathbf{Top}$ be topological spaces. Then the *product space* $A \times B$ is a topological space with the Cartesian product as the underlying set of points, turned into a topological space via the **product topology**. This is the topology whose opens are unions of sets of the form $U \times V$, where $U \subseteq A$ and $V \subseteq B$ are open sets in the respective spaces. Then the projections

$$p_A : A \times B \longrightarrow A, \quad p_B : A \times B \longrightarrow B,$$

are continuous, since sets of the form $U \times B$ and $A \times V$ are open in $A \times B$ by definition.

Let us verify that the universal property of the categorical product is satisfied. If we have another topological space X and two continuous maps $f : X \rightarrow A$ and $g : X \rightarrow B$, then the universal property of the product in **Set** gives us the unique map $(f, g) : X \rightarrow A \times B$ satisfying the relevant equations $p_A \circ (f, g) = f$ and $p_B \circ (f, g) = g$. But is (f, g) really continuous? To see this, it is enough to check it on the basis of opens, meaning on the open sets of the form $U \times V$. But there, we have

$$\begin{aligned} (f, g)^{-1}(U \times V) &= \{x \in X \mid (f(x), g(x)) \in U \times V\} \\ &= \{x \in X \mid f(x) \in U \text{ and } g(x) \in V\} \\ &= f^{-1}(U) \cap g^{-1}(V). \end{aligned}$$

Since $f^{-1}(U)$ and $g^{-1}(V)$ are open in X , so is their intersection, and we conclude that (f, g) is continuous.

Typically, there will be many other topologies which make the projections p_A and p_B continuous. But the product topology is singled out by the universal property by the essential uniqueness of the universal object: any other categorical product space must be homeomorphic to $A \times B$ such that the homeomorphism respects the projections, and therefore the open sets must precisely correspond.

Example 4.1.5 (Cartesian product of groups). Do categorical products exist in **Grp**? Yes, they do, and it is probably not surprising at this point that they are given by the usual Cartesian product of groups, also known as the **direct product**.

So for two groups G and H , their Cartesian product $G \times H$ becomes a group via the pointwise multiplication,

$$(g_1, h_1) \cdot (g_2, h_2) := (g_1 g_2, h_1 h_2).$$

It is easy to see that this makes the projections $p_G : G \times H \rightarrow G$ and $p_H : G \times H \rightarrow H$ into group homomorphisms. In contrast to the case of the product topology, where in general many other choices of topology can achieve this, this is the *only* group structure on $G \times H$ which makes the projections into group homomorphisms.

Now if X is any group and $f : X \rightarrow G$ and $g : X \rightarrow H$ are group homomorphisms, then the universal property of the categorical product in **Set** gives us the pairing $(f, g) : X \rightarrow G \times H$. Since

$$\begin{aligned} (f, g)(x_1 x_2) &= (f(x_1 x_2), g(x_1 x_2)) \\ &= (f(x_1) f(x_2), g(x_1) g(x_2)) \\ &= (f(x_1), g(x_1)) \cdot (f(x_2), g(x_2)) \\ &= (f, g)(x_1) \cdot (f, g)(x_2), \end{aligned}$$

this pairing (f, g) is a group homomorphism too, and we have shown that the universal property of the categorical product is satisfied.

A bit of thought shows that the previous example is not very specific to group theory: it works the same way for many other kinds of algebraic structures like rings, monoids, modules or Lie algebras, where one obtains the categorical product by taking the Cartesian product and defining all algebraic operations componentwise. However, we need to be careful that the algebraic structures under consideration are defined only by “operations plus laws”; the following example shows that imposing *non-equational conditions* in the definition of the algebraic structure can mess up the situation.

Example 4.1.6. Let $\mathbf{Fld}$ be the category of fields, by which we mean that the objects are fields and the morphisms are ring homomorphisms. Then $\mathbf{Fld}$ does not have a categorical product of $\mathbb{C}$ with itself, as we now show.¹⁸ To see this, suppose that K is a categorical product of $\mathbb{C}$ with itself with respect to projections

$$p_1 : K \longrightarrow \mathbb{C}, \quad p_2 : K \longrightarrow \mathbb{C}.$$

We then apply the universal property to $\text{id}_{\mathbb{C}}$ and itself, which results in the paired morphism¹⁹

$$(\text{id}_{\mathbb{C}}, \text{id}_{\mathbb{C}}) : \mathbb{C} \longrightarrow K.$$

The equation $p_1 \circ (\text{id}_{\mathbb{C}}, \text{id}_{\mathbb{C}}) = \text{id}_{\mathbb{C}}$ shows that p_1 is surjective. Since a ring homomorphism between fields is always injective, we can in fact conclude that p_1 is bijective, and therefore an isomorphism. Consider now also the pairing $(\text{id}_{\mathbb{C}}, cc)$, where $cc : \mathbb{C} \rightarrow \mathbb{C}$ is complex conjugation. Then we have

$$cc = p_2 \circ (\text{id}_{\mathbb{C}}, cc) = p_2 \circ p_1^{-1} \circ \text{id}_{\mathbb{C}} = p_2 \circ (\text{id}_{\mathbb{C}}, \text{id}_{\mathbb{C}}) = \text{id}_{\mathbb{C}},$$

where the second and third step use $p_1 \circ (\text{id}_{\mathbb{C}}, cc) = \text{id}_{\mathbb{C}} = p_1 \circ (\text{id}_{\mathbb{C}}, \text{id}_{\mathbb{C}})$. Clearly $cc = \text{id}_{\mathbb{C}}$ is not true, and we thus have arrived at a contradiction and can conclude that $\mathbb{C}$ does not have a categorical product with itself in $\mathbf{Fld}$.

Counterintuitively, *some* categorical products do exist in $\mathbf{Fld}$, although they are not given by the Cartesian product with its usual projections. For example, the categorical product of $\mathbb{Q}$ with itself is given *by $\mathbb{Q}$ itself* again, with both product projections $\mathbb{Q} \rightarrow \mathbb{Q}$ being the identity morphism. Indeed the universal property can be derived easily from the following observations, which apply to any field K :

- ▷ If $K \cong \mathbb{Q}$, then there is precisely one homomorphism $K \rightarrow \mathbb{Q}$.
- ▷ Otherwise, there is no homomorphism $K \rightarrow \mathbb{Q}$ at all.

The first observation above holds because the identity map is the only morphism $\mathbb{Q} \rightarrow \mathbb{Q}$, and we can compose with the inverse of $K \cong \mathbb{Q}$ in order to obtain the desired statement. For the second, we simply recall that every homomorphism of fields is injective, so that $K \rightarrow \mathbb{Q}$ would make K into a subfield of $\mathbb{Q}$, but the only subfield of $\mathbb{Q}$ is itself, which gives $K \cong \mathbb{Q}$.

By the second observation, we only need to check the universal property on pairs of morphisms out of $K = \mathbb{Q}$. But in this case, the only possibility is the identity morphism. Hence there is only one single instance of the universal property, namely the pairing $(\text{id}_{\mathbb{Q}}, \text{id}_{\mathbb{Q}})$, which coincides with $\text{id}_{\mathbb{Q}}$ again.

Example 4.1.7 (Meets in a preordered set). We saw in Example 2.1.7 that the (small) categories in which all hom-sets are either empty or have exactly one element can be identified with preordered sets, where one interprets the existence of a morphism $a \rightarrow b$ as the statement $a \leq b$.

Now what is the categorical product of two objects a, b in a preordered set, and when does it exist? By definition, a categorical product amounts to an element $a \times b$, which we now write as $a \wedge b$ instead, such that the following hold:

- ▷ $a \wedge b \leq a$ and $a \wedge b \leq b$.
- ▷ For every element x with $x \leq a$ and $x \leq b$, we have $x \leq a \wedge b$.

¹⁸If you prefer, you can also do the same argument with the Gaussian rationals $\mathbb{Q}(i)$ instead of $\mathbb{C}$, or with $\mathbb{Q}(\sqrt{2})$, or really with any field admitting a nontrivial automorphism playing the role of cc .

¹⁹In any category, if A is an object for which the product $A \times A$ exists, then applying the universal property to $\text{id}_A : A \rightarrow A$ and itself gives a morphism $A \rightarrow A \times A$ called the **diagonal** of A . We consider this here in $\mathbf{Fld}$ for $A = \mathbb{C}$.

Note that the uniqueness part of the universal property is trivial since any two morphisms between the same two objects are trivially equal. Therefore we only need to care about the existence part, which manifests itself as the inequalities above. Such an element $a \wedge b$ is called the **meet** of a and b in the preordered set.

The uniqueness of the universal object $a \wedge b$ now means that for any two choices of meet, each is $\leq$ than the other. This makes them indistinguishable from the point of view of the preorder relation $\leq$, and therefore we speak of *the* meet of a and b .

The universal property will perhaps be most familiar to you in the case of the integers $\mathbb{Z}$, if we let $a \leq b$ denote “ a divides b ”. Then the meet of $a, b \in \mathbb{Z}$ is called the *greatest common divisor* (*gcd*) of a and b , as the first item above states that $a \wedge b$ divides both a and b , and the second states that any common divisor of a and b divides $a \wedge b$. The uniqueness now corresponds to the fact that the gcd is uniquely determined up to sign, since two integers, where each divides the other, are equal up to sign.²⁰

More elementary are meets in $\mathbb{R}$, considered as a preordered set with the usual ordering $\leq$. There, the meet of any two numbers $a, b \in \mathbb{R}$ is the *minimum* $\min(a, b)$, meaning simply the smaller of the two. Because we indeed have $\min(a, b) \leq a$ and $\min(a, b) \leq b$, and if $x \leq a$ and $x \leq b$, then also $x \leq \min(a, b)$.

It seems worth noting the following direct consequence of the universal property separately.

Lemma 4.1.8. *Suppose that objects A and B have a product as above.*

(a) *For every $f : X \rightarrow A \times B$, we have*

$$(p_A f, p_B f) = f.$$

In particular, taking $f = \text{id}_{A \times B}$ shows $(p_A, p_B) = \text{id}_{A \times B}$.

(b) *For every $f, g : X \rightarrow A \times B$, we have*

$$f = g \quad \Longleftrightarrow \quad p_A f = p_A g \quad \wedge \quad p_B f = p_B g.$$

Proof. (a) By the universal property, the pairing $(p_A f, p_B f)$ is the unique morphism whose composition with p_A is $p_A f$ and whose composition with p_B is $p_B f$. But since f displays these properties as well, we conclude that $(p_A f, p_B f) = f$ by the uniqueness part of the universal property.

(b) This can either be shown using the same idea, or by the formula just proven,

$$f = (p_A f, p_B f) = (p_A g, p_B g) = g. \quad \square$$

To what extent is a categorical product unique? For universal objects in general, we know that they are “unique up to unique isomorphism” by Proposition 3.3.6. In the setting of products, unfolding the definitions shows that this turns into the following statement.

Lemma 4.1.9. *If $A \times B$ and $A \otimes B$ are two products of A and B with projections p_A, p_B and q_A, q_B , respectively, then there is a unique isomorphism*

$$\phi : A \times B \xrightarrow{\cong} A \otimes B$$

which is compatible with the projections in the sense that

$$q_A \phi = p_A, \quad q_B \phi = p_B. \quad (4.2)$$

²⁰All of this applies much more generally to divisibility of elements in more general commutative rings, namely unique factorization domains, but this need not concern us here.

These statements can be used in many proofs, as we will see in the following.

Proof. Instead of obtaining this as a special case of Proposition 3.3.6, we can also argue directly like this. The universal property of $A \otimes B$ gives us a morphism $\phi : A \times B \rightarrow A \otimes B$ such that (4.2) holds. Similarly, by the universal property of $A \times B$, we obtain a morphism $\psi : A \otimes B \rightarrow A \times B$ such that

$$p_A \psi = q_A, \quad p_B \psi = q_B.$$

It remains to be shown that ψ is the inverse of ϕ . Indeed we get $\phi\psi = \text{id}_{A \otimes B}$ by Lemma 4.1.8, since composing both sides with q_A gives q_A again, and similarly for q_B . The proof of $\psi\phi = \text{id}_{A \times B}$ is analogous. $\square$

As category theorists, we often want to ensure that a given notion or constructions is functorial. In the case of products, this means: if objects A and B have a product $A \times B$, and C and D have a product $C \times D$, then do we obtain a canonical morphism $A \times B \rightarrow C \times D$ from given morphisms $f : A \rightarrow C$ and $g : B \rightarrow D$?

To see that this can indeed be done, consider the diagram:²¹

$$\begin{array}{ccccc} & & A & \xrightarrow{f} & C \\ & \nearrow p_A & & & \nearrow p_C \\ A \times B & \dashrightarrow (fp_A, gp_B) & C \times D & & \\ & \searrow p_B & & & \searrow p_D \\ & & B & \xrightarrow{g} & D \end{array} \quad (4.3)$$

In **Set** and similar categories like **Top** and **Grp**, the morphism $f \times g : A \times B \rightarrow C \times D$ from (4.3) is exactly given by what you'd expect, namely applying f to the first component and g to the second component. This holds because the first component of the result is fp_A , which takes a pair (x, y) and first projects to x before applying f , and similarly the second component gp_B projects to y and then applies g .

In order to formally formulate the functoriality, we consider $\mathcal{C}^2$, the category where objects are pairs of objects (A, B) from $\mathcal{C}$, and morphisms $(A, B) \rightarrow (C, D)$ are pairs of morphisms $(f : A \rightarrow C, g : B \rightarrow D)$. Note that this can be viewed as a functor category $[\mathbf{2}, \mathcal{C}]$, where $\mathbf{2}$ is the category with two objects and only identity morphisms.

Lemma 4.1.10. *Let $\mathcal{C}$ be a category together with a choice of product for any two objects.²² Then with the action on morphisms defined as in (4.3), the formation of products is a functor $\mathcal{C}^2 \rightarrow \mathcal{C}$.*

We also denote the dashed morphism in (4.3) by $f \times g$. Like this, the operation “ $\times$ ” can be viewed as a functor into which one can plug in either a pair of objects or a pair of morphisms.

Proof. We only need to show that the construction of (4.3) preserves composition and identities. The preservation of identities amounts to the equation $(p_A, p_B) = \text{id}_{A \times B}$, which we have proven in Lemma 4.1.8.

The preservation of composition follows similarly from the uniqueness part of the universal property. To see this in detail, let (E, F) be a third pair of objects, and let $h : C \rightarrow E$ and

²¹Note that we have already used something very similar in the proof of Lemma 4.1.9, where we had $f = \text{id}_A$ and $g = \text{id}_A$ with two choices of product $A \times A$ and $A \otimes A$.

²²This concludes the assumption that any two objects *have* a product to start with.

$k : D \rightarrow F$ be morphisms. Then we need to consider the diagram

$$\begin{array}{ccccc}
 & A & \xrightarrow{f} & C & \xrightarrow{h} & E \\
 & \nearrow p_A & \searrow (hf p_A, k g p_B) & \nearrow p_C & \searrow p_E & \\
 A \times B & \xrightarrow{(f p_A, g p_B)} & C \times D & \xrightarrow{(h p_C, k p_D)} & E \times F & \\
 & \searrow p_B & \nearrow p_D & \searrow p_F & & \\
 & B & \xrightarrow{g} & D & \xrightarrow{k} & F
 \end{array}$$

and prove that the composite of the two dashed morphisms is equal to the dotted morphism. To this end, by the universal property it is enough to prove this after composing with p_E and with p_F by Lemma 4.1.8. But this follows from the commutativity of the diagram by a straightforward diagram chase. $\square$

For a category with a terminal object 1 , we write $!_A : A \rightarrow 1$ for the unique morphism from A to 1 .

Lemma 4.1.11. *Suppose that a category has a terminal object 1 . Then every object A serves as a product $A \times 1$ with respect to the projections*

$$\text{id}_A : A \rightarrow A, \quad !_A : A \rightarrow 1.$$

Proof. The only morphism $g : X \rightarrow 1$ is the unique morphism $!_X$. Using this, it is straightforward to see that the unique choice for the pairing of $f : X \rightarrow A$ and $!_X : X \rightarrow 1$ is $(f, !_X) := f$ itself. $\square$

Lemma 4.1.12. *Suppose that in a category $\mathcal{C}$, we have chosen a product $A \times B$ for every pair of objects $A, B \in \mathcal{C}$ and a terminal object 1 . Then there are natural isomorphisms*

$$1 \times A \xrightarrow{\cong} A \xleftarrow{\cong} A \times 1.$$

Here, the naturality means that for any morphism $f : A \rightarrow A'$, the following diagram commutes:

$$\begin{array}{ccc}
 1 \times A & \xrightarrow{\cong} & A \\
 \text{id}_1 \times f \downarrow & & \downarrow f \\
 1 \times A' & \xrightarrow{\cong} & A'
 \end{array} \tag{4.4}$$

These are two naturality squares if we consider both the left and the right as functors $\mathcal{C} \rightarrow \mathcal{C}$, where the vertical arrows denote the action of the functor on a morphism f , so that in particular the second functor is the identity functor.

Proof. The product projection $1 \times A \rightarrow A$ is an isomorphism, with inverse given by the pairing $(!_A, \text{id}_A)$. The fact that these two are each other's inverses follows e.g. by an application of Lemma 4.1.8.

The commutativity of the naturality square (4.4) then amounts to the statement that the second component of $\text{id}_1 \times f$ is precisely the projection followed by f , and this holds by the construction of $\text{id}_1 \times f$. $\square$

What we have considered so far are **binary products**, i.e. categorical products of exactly two objects. Bu the definition generalizes easily to a definition of product for an arbitrary family $(A_i)_{i \in I}$ of objects, where I is any (finite or infinite) index set.

Definition 4.1.13. Let $(A_i)_{i \in I}$ be a family of objects in a category $\mathcal{C}$. Then the **categorical product** $\prod_{i \in I} A_i$ is an object $A := \prod_{i \in I} A_i$ together with **product projections**

$$p_j : \prod_{i \in I} A_i \longrightarrow A_j$$

such that for every family of morphisms $(f_i : X \rightarrow A_i)_{i \in I}$, there is a unique morphism $f : X \rightarrow \prod_{i \in I} A_i$ such that

$$f_i = p_j f \quad \forall i \in I. \quad (4.5)$$

In other words, a product is a representing object for the functor $\mathcal{C}^{\text{op}} \rightarrow \mathbf{Set}$ which assigns to every object X the set of all families of morphisms $(f_i : X \rightarrow A_i)_{i \in I}$, and operates on morphisms by composition as usual.

When I has precisely two elements, then we recover the definition of the binary product from Definition 4.1.1. It is also possible to consider $I = \emptyset$. In this case, for every X the family (f_i) is unique (namely the empty family), and the factorization property (4.5) holds vacuously. Thus the universal property is that of a terminal object: the empty product exists in $\mathcal{C}$ if and only if $\mathcal{C}$ has a terminal object. This is analogous to how an empty product of numbers is defined to be 1, and explains why we usually denote a terminal object also by 1.

Definition 4.1.14. A category $\mathcal{C}$ **has finite products** if a product $\prod_{i \in I} A_i$ exists for every family $(A_i)_{i \in I}$ of objects in $\mathcal{C}$ with I finite.

Similar to how forming products of more than two numbers can be done by repeated application of the binary product, we can likewise reduce finite products to the binary (and nullary) case, as follows.

Lemma 4.1.15. A category $\mathcal{C}$ has finite products if and only if it has binary products and a terminal object.

Proof. The “only if” direction is clear, since we have already seen that the existence of binary products and of a terminal object are special cases of the existence of finite products.

For the “if” direction, let $(A_i)_{i \in I}$ be a finite family of objects in $\mathcal{C}$, and suppose that the index set is $I = \{1, \dots, n\}$ without loss of generality. Then we construct a product $\prod_{i=1}^n A_i$ by induction on n . For $n = 0$, we can take the terminal object 1 as the product. For $n = 1$, we have a single object A_1 , and the product is just A_1 itself, with the identity morphism as the product projection. For $n = 2$, we are in the situation of the binary product.

So it remains to perform the induction step for $n \geq 3$, for which we will argue that

$$\prod_{i=1}^n A_i := \left(\prod_{i=1}^{n-1} A_i \right) \times A_n \quad (4.6)$$

works with a suitable choice of product projections. If p_1 and p_2 are the product projections of the binary product and $q_1, \dots, q_{n-1}$ those of the product $\prod_{i=1}^{n-1} A_i$, then we construct new product projections as

$$r_1 := q_1 p_1, \quad \dots, \quad r_{n-1} := q_{n-1} p_1, \quad r_n := p_2.$$

To prove the universal property for these, suppose that we are given morphisms $f_i : X \rightarrow A_i$ for $i = 1, \dots, n$. Then the universal property of $\prod_{i=1}^{n-1} A_i$ gives us a unique morphism

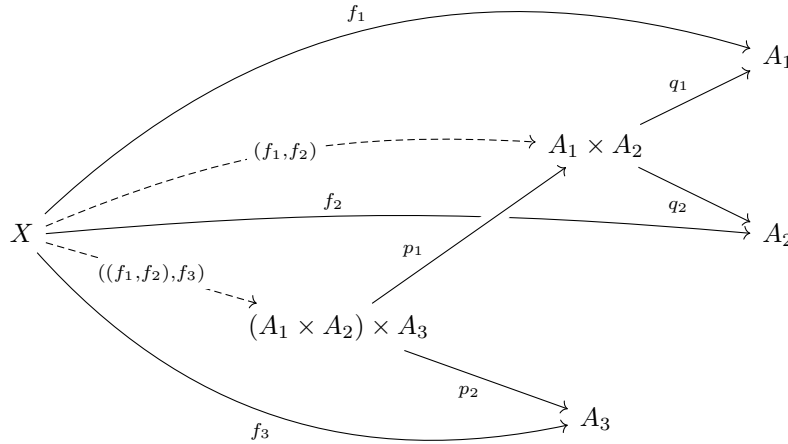
$$(f_1, \dots, f_{n-1}) : X \longrightarrow \prod_{i=1}^{n-1} A_i$$

with the property that $q_i \circ (f_1, \dots, f_{n-1}) = f_i$ for $i = 1, \dots, n-1$. Applying then the universal property of the binary product with projections p_1 and p_2 then gives us a unique $f : X \rightarrow \prod_{i=1}^n A_i$ such that

$$p_1 f = (f_1, \dots, f_{n-1}) \quad \text{and} \quad p_2 f = f_n.$$

Composing with the q_i then gives us the desired equations $r_i f = f_i$ for $i = 1, \dots, n$. The uniqueness follows from the uniqueness of the morphisms in both applications of the universal property. $\square$

It may help to illustrate the proof for $n = 3$ by the following commutative diagram:



Remark 4.1.16. By working with $A_1 \times (A_2 \times A_3)$ instead of $(A_1 \times A_2) \times A_3$, we can similarly show that the former also serves as a product of A_1, A_2, A_3 . In particular, we can thus obtain an isomorphism

$$A_1 \times (A_2 \times A_3) \xrightarrow{\cong} (A_1 \times A_2) \times A_3.$$

This is furthermore natural in the sense that for any morphisms $f : A_1 \rightarrow A'_1$, $g : A_2 \rightarrow A'_2$ and $h : A_3 \rightarrow A'_3$, the following diagram commutes:

$$\begin{array}{ccc} A_1 \times (A_2 \times A_3) & \xrightarrow{\cong} & (A_1 \times A_2) \times A_3 \\ f \times (g \times h) \downarrow & & \downarrow (f \times g) \times h \\ A'_1 \times (A'_2 \times A'_3) & \xrightarrow{\cong} & (A'_1 \times A'_2) \times A'_3. \end{array} \quad (4.7)$$

This naturality can be derived by using the projections constructed in the proof of Lemma 4.1.15 and applying Lemma 4.1.9.

There are other natural isomorphisms which can be constructed by similar reasoning, such as

$$A \times B \cong B \times A.$$

Example 4.1.17. By what we already know, we can conclude that **Set**, **Top** and **Grp** all have finite products. (In fact they have products for any family of objects, but we will not need this.)

On the other hand, **Fld** does not have finite products, as we have seen in Definition 4.1.1. And already because it does not have a terminal object!

Remark 4.1.18. Before learning category theory, how would you have answered the question of how a ternary Cartesian product of sets $A \times B \times C$ is defined? Is it $(A \times B) \times C$, or $A \times (B \times C)$, or something else? There does not seem to be an answer to this that would be both precise and not arbitrary.

But in category theory, we have a precise answer: $A \times B \times C$ stands for *any* object which satisfies the universal property of a product, and the uniqueness of the universal object guarantees that the particular choice is completely irrelevant for mathematical practice. Moreover, arguing as in the proof of Lemma 4.1.15 shows that both $(A \times B) \times C$ and $A \times (B \times C)$ can be used as the product of

the three objects, and they are in fact naturally²³ isomorphic to each other. In other words, binary products are associative up to natural isomorphism.

Finite products are particularly useful for making sense of algebraic structures like groups in *any* category with finite products. We can do this by simply writing down the standard definition phrased in terms of maps and Cartesian products rather than elements, and then generalizing the Cartesian products to categorical products.

Definition 4.1.19. *Let $\mathcal{C}$ be a category with finite products. Then a **group internal to $\mathcal{C}$** is an object G together with morphisms*

$$m : G \times G \longrightarrow G, \quad e : 1 \longrightarrow G, \quad i : G \longrightarrow G$$

such that the following diagrams commute:

▷ *Associativity:*

$$\begin{array}{ccc} (G \times G) \times G & \xrightarrow{\cong} & G \times (G \times G) \\ m \times \text{id}_G \downarrow & & \downarrow \text{id}_G \times m \\ G \times G & & G \times G \\ & \searrow m & \swarrow m \\ & G & \end{array} \quad (4.8)$$

▷ *Unit laws:*

$$\begin{array}{ccccc} 1 \times G & \xrightarrow{e \times \text{id}_G} & G \times G & \xleftarrow{\text{id}_G \times e} & G \times 1 \\ & \searrow \cong & \downarrow m & \swarrow \cong & \\ & & G & & \end{array} \quad (4.9)$$

▷ *Inverse laws:*

$$\begin{array}{ccccc} & & G & & \\ & \swarrow (\text{id}_G, i) & \downarrow !_G & \searrow (i, \text{id}_G) & \\ G \times G & & 1 & & G \times G \\ & \searrow m & \downarrow e & \swarrow m & \\ & & G & & \end{array} \quad (4.10)$$

A group internal to **Set** is just a group in the usual sense, as one can see by writing the commutativity of the diagrams in terms of elements while using the fact that the categorical products are Cartesian products. When you do this, you'll find that these conditions turn into the following, which are required to hold for all elements $x, y, z \in G$:

▷ *Associativity:* $m(m(x, y), z) = m(x, m(y, z))$.

▷ *Unit laws:* $m(e, x) = x = m(x, e)$.

▷ *Inverse laws:* $m(x, i(x)) = e = m(i(x), x)$.

²³Puzzle: what is the precise meaning of “naturally” here, and why does this naturality hold?

A group internal to **Top** is a **topological group**, meaning a group that is also a topological space such that the multiplication and inversion maps are continuous, where for the multiplication one uses the product topology on $G \times G$. Topological groups are important in many areas of mathematics and theoretical physics and have a rich theory of their own. It is a good sign that our abstract definition reproduces a commonly studied concept!

4.2 Equations and equalizers

Especially from the perspective of a layperson, mathematics is all about solving equations. But what does “solving” an equation actually mean? Usually, it means finding a description of the set of *solutions* to the equation which is as explicit as possible.

An equation

$$f(x) = g(x) \quad (4.11)$$

involves two functions f and g , defined on a common set X that the variable x ranges over. In order for the equation to make sense, the functions f and g also must have the same codomain Y . The set of solutions is then a subset $E \subseteq X$. If we write $i : E \hookrightarrow X$ for the inclusion map, then we obtain the diagram

$$E \xrightarrow{i} X \begin{array}{c} \xrightarrow{f} \\ \xrightarrow{g} \end{array} Y \quad (4.12)$$

where $fi = gi$. This diagram is *not* commutative, because commutativity would require $f = g$, in which case the situation trivializes: an equation (4.11) in which both sides already are the same function is not interesting, because its set of solutions is just X itself.

Now in order for E to count as the set of solutions, we also want it to be *maximal* with the property that f and g agree on it. This is precisely the content of the following universal property:

Definition 4.2.1. *Given morphisms $f, g : X \rightarrow Y$ in a category $\mathcal{C}$, an **equalizer** of f and g is an object E together with a morphism $i : E \rightarrow X$ such that:*

- (a) $fi = gi$.
- (b) *For every $A \in \mathcal{C}$ and $h : A \rightarrow X$ with $fh = gh$, there is a unique morphism $\hat{h}$ such that $h = i\hat{h}$.*

This universal property is best illustrated by the diagram

$$\begin{array}{ccccc} & & A & & \\ & \swarrow \hat{h} & & \searrow h & \\ E & \xrightarrow{i} & X & \begin{array}{c} \xrightarrow{f} \\ \xrightarrow{g} \end{array} & Y \end{array}$$

Now let us consider this in the context of our equation (4.11) again, which takes place in **Set**. Then with E being the set of solutions to the equation and $i : E \hookrightarrow X$ the inclusion map, the universal property indeed holds: the assumption $fh = gh$ means that $h(a)$ is a solution to the equation for every $a \in A$, and hence it factors across E . This defines the map $\hat{h} : A \rightarrow E$, and the equation $h = i\hat{h}$ holds by construction. The uniqueness of $\hat{h}$ is obvious because i is a monomorphism.

Equalizers exist and behave similarly as in **Set** in many categories.

Example 4.2.2. In **Top**, the equalizer of two continuous maps $f, g : X \rightarrow Y$ is likewise given by the solution set E with the inclusion map $i : E \hookrightarrow X$. But now we also need to specify a topology on E . The only topology which works is the **subspace topology**, where the open sets are precisely the sets of the form $i^{-1}(U) = E \cap U$ for open sets $U \subseteq X$.

Indeed to check the universal property as above, we now only need to verify that $\hat{h}$ is continuous if h is. But this is clear: for an open set of the form $i^{-1}(U)$ as above, we have

$$\hat{h}^{-1}(i^{-1}(U)) = (i\hat{h})^{-1}(U) = h^{-1}(U),$$

which is open by assumption.

Example 4.2.3. In **Grp**, the equalizer of two group homomorphisms $f, g : X \rightarrow Y$ is given by the solution set E with the inclusion map $i : E \hookrightarrow X$. Since i must become a group homomorphism, the only way to define the group structure on E is to restrict the one from X , i.e. we need to ensure that E is a subgroup of X . But this is straightforward: if $x_1, x_2 \in E$ are two solutions to the equation $f(x) = g(x)$, then so is their product x_1x_2 by

$$f(x_1x_2) = f(x_1)f(x_2) = g(x_1)g(x_2) = g(x_1x_2),$$

using the assumption that f and g are group homomorphisms. Similarly, E contains the neutral element e , because of

$$f(e) = e = g(e).$$

The universal property now holds because it holds in **Set**, and $\hat{h}$ is easily seen to be a group homomorphism as well.

Equalizers in general thus provide us with an extremely general definition of what it means to solve an equation. We can even prove at this level of generality that the solution object E is “included” in X , which formally means that $i : E \hookrightarrow X$ is a monomorphism.

Lemma 4.2.4. *Let*

$$E \xrightarrow{i} X \rightrightarrows^f_g Y$$

be an equalizer diagram in any category $\mathcal{C}$. Then i is a monomorphism.

Proof. Let $k_1, k_2 : A \rightarrow E$ be two morphisms such that $ik_1 = ik_2$, and let us write $h : A \rightarrow X$ for this composite, meaning

$$h := ik_1 = ik_2,$$

as illustrated in the partially commutative diagram

$$\begin{array}{ccccc} A & & & & \\ & \searrow^{k_2} & & \searrow^h & \\ & & E & \xrightarrow{i} & X \rightrightarrows^f_g Y \\ & \swarrow_{k_1} & & & \end{array}$$

Then we have $fh = gh$ because of²⁴

$$fh = fik_1 = gik_1 = gh.$$

The universal property gives us a *unique* $\hat{h} : A \rightarrow X$ such that $h = i\hat{h}$. Since both k_1 and k_2 satisfy this in place of $\hat{h}$, the uniqueness requires $k_1 = k_2$. $\square$

²⁴Note that we could just as well use k_2 instead of k_1 here.

4.3 Limits in general

Both products and equalizers are mapping-in universal properties. They are of a similar flavour: both start with a diagram in $\mathcal{C}$ and then consider how to “map into” this diagram. (In the case of products, the original diagram just consists of two objects with no morphisms between them.) This indicates that the two notions might have a common generalization, and we develop it now.

Recall from Definition 2.8.1 that a **diagram** in a category $\mathcal{C}$ is given by a direct graph $G = (V, E, s, t)$ and labelling functions

$$\ell_V : V \longrightarrow \text{Ob}(\mathcal{C}), \quad \ell_E : E \longrightarrow \text{Mor}(\mathcal{C}).$$

To have a simpler notation, we will now omit the subscripts and just write ℓ for both maps, as one can tell which one is meant by $\ell(x)$ depending on whether x is a vertex or an edge.

Definition 4.3.1. (a) Given such a diagram in $\mathcal{C}$, a **cone** is a family of morphisms $(f_v : X \rightarrow \ell(v))_{v \in V}$ for some object X such that

$$f_{t(e)} = \ell(e) f_{s(e)} \quad \forall e \in E. \quad (4.13)$$

(b) A cone $(u_v : L \rightarrow \ell(v))_{v \in V}$ is a **limit cone** if for every cone $(f_v : X \rightarrow \ell(v))_{v \in V}$ there is a unique morphism $g : X \rightarrow L$ such that

$$f_v = u_v g \quad \forall v \in V,$$

and the object L is called the **limit** of the diagram.

For a cone $(f_v : X \rightarrow \ell(v))_{v \in V}$, we call X the **apex** of the cone, and the morphisms f_v the **legs**. To understand the definition of limit as a universal property, we consider the functor which assigns to every object X the set of cones with apex X and acts on morphisms by composition,

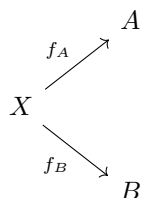
$$\begin{aligned} \mathcal{C}^{\text{op}} &\longrightarrow \text{Set} \\ X &\longmapsto \{\text{cones } (f_v : X \rightarrow \ell(v))_{v \in V}\} \\ (g : X \rightarrow Y) &\longmapsto ((f_v : Y \rightarrow \ell(v))_{v \in V} \mapsto (f_v g : X \rightarrow \ell(v))_{v \in V}). \end{aligned}$$

Then we can rephrase the definition of limit cone as saying that it’s a universal element of this functor. By Proposition 3.3.1, this makes the limit L into a representing object for the functor.

Example 4.3.2. (a) Consider first the graph consisting only of two isolated vertices:

$$\bullet \qquad \bullet \quad (4.14)$$

Then a diagram of this shape in a category $\mathcal{C}$ is simply a pair of objects A and B , and a cone with apex X is a diagram like this:



By Definition 4.1.1, a limit is precisely a product, where the legs of the limit cone are the product projections $p_A : L \rightarrow A$. So for the graph (4.14), the definition of limit specializes to the definition of product.

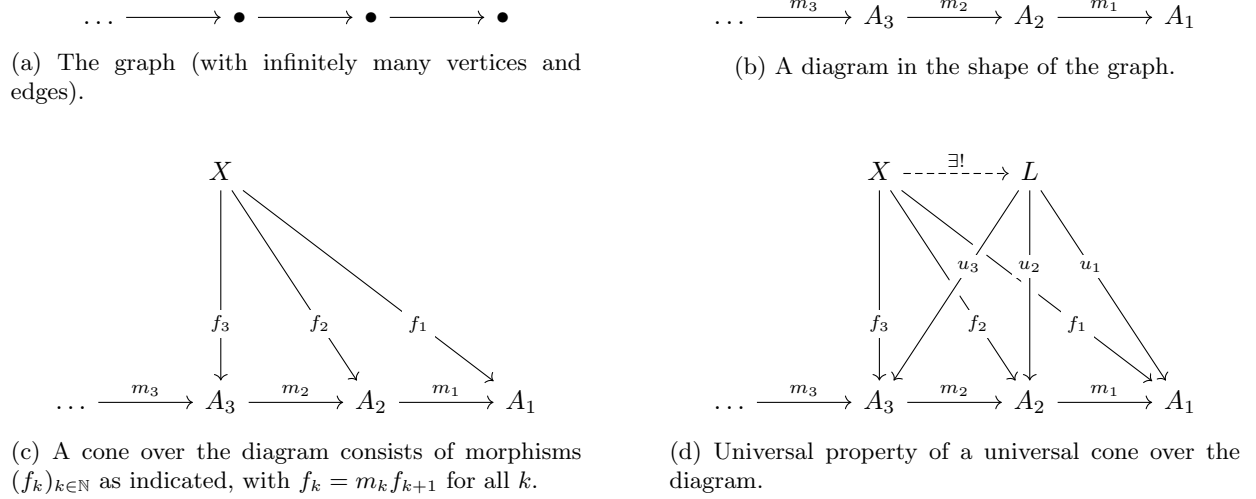


Figure 4: Illustration of the universal property of a projective limit.

(b) Similarly, for the graph

$$\bullet \rightrightarrows \bullet$$

the definition of limit specializes to the definition of equalizer. Indeed a diagram of this shape looks like

$$A \begin{smallmatrix} \xrightarrow{g} \\ \xrightarrow{h} \end{smallmatrix} B$$

as considered before, and a cone with apex X comes with additional morphisms like this

$$\begin{array}{ccc} X & & \\ & \searrow f_A & \searrow f_B \\ & A & \xrightarrow[k]{h} B \end{array}$$

where the extra conditions (4.13) now read $f_B = hf_A$ and $f_B = kf_A$, which means in particular that $hf_A = kf_A$. Since f_B is uniquely determined, a cone is uniquely specified by giving $f_A : X \rightarrow A$ such that $hf_A = kf_A$. Based on this observation, we can see that the universal property of a limit cone is precisely the one of an equalizer as in Definition 4.2.1 (now with different notation): a limit cone with components u_A and u_B also satisfies $u_B = hu_A = ku_A$, and the universal factorization gives us a unique $\hat{f} : X \rightarrow L$ such that

$$f_A = u_A \hat{f} \quad \text{and} \quad f_B = u_B \hat{f}.$$

But since $f_B = hf_A$ and $u_B = hu_A$, the second condition is a consequence of the first one obtained by composing it with h , and therefore the second condition can be omitted. This reproduces exactly the definition of equalizer.

Figure 4(c) illustrates the concept of limit for another shape of diagram, for which a limit is known as **projective limit**.²⁵ Formally, we can define the graph as

$$V := \mathbb{N}, \quad E := \mathbb{N}, \quad s(k) := k + 1, \quad t(k) := k,$$

²⁵Sometimes this is called **inverse limit**, but we avoid this term in order to not confuse it with inverses of morphisms.

meaning that both vertices and edges are enumerated by natural numbers, and the n -th edge goes from the vertex $n + 1$ to the vertex n .

Working with projective limits for a bit will help familiarize ourselves more with the concept of limit in general. First of all, specializing the general definition of limit to the case of projective limit gives us the following definition:

Definition 4.3.3. Let $(A_k)_{k \in \mathbb{N}}$ be a sequence of objects in a category $\mathcal{C}$ and

$$(m_k : A_{k+1} \longrightarrow A_k)_{k \in \mathbb{N}}$$

a sequence of morphisms. Then a **cone** is an object X together with morphisms

$$f_k : X \longrightarrow A_k$$

such that for every $k \in \mathbb{N}$, the diagram

$$\begin{array}{ccc} & X & \\ f_{k+1} \swarrow & & \searrow f_k \\ A_{k+1} & \xrightarrow{m_k} & A_k \end{array}$$

commutes. A **projective limit** is a cone $(u_k : L \rightarrow A_k)_{k \in \mathbb{N}}$ enjoying the unique factorization property: for every cone $(f_k : X \rightarrow A_k)_{k \in \mathbb{N}}$, there is a unique $g : X \rightarrow L$ with $f_k = u_k g$ for all $k \in \mathbb{N}$.

It is interesting to note that for such a cone, each morphism f_k is uniquely determined by f_{k+1} , which in turn is uniquely determined by f_{k+2} , and so on. In this sense, a cone consists of a sequence of successive “liftings” of each f_k to the next f_{k+1} .

Example 4.3.4. In **Top**, let us consider the unit interval $[0, 1]$, and its powers $[0, 1]^n$ for $n \in \mathbb{N}$ equipped with the product topology. Then for $n \geq 1$, we can identify $[0, 1]^n$ with the binary product $[0, 1]^{n-1} \times [0, 1]$ as in the proof of Lemma 4.1.15. The corresponding product projections are $p_n : [0, 1]^n \rightarrow [0, 1]^{n-1}$, and they are the maps which simply project onto the first $n - 1$ coordinates. These are the maps in the diagram

$$\dots \xrightarrow{p_4} [0, 1]^4 \xrightarrow{p_3} [0, 1]^3 \xrightarrow{p_2} [0, 1]^2 \xrightarrow{p_1} [0, 1].$$

in **Top**. The limit of this diagram exists and is given by the infinite product $[0, 1]^{\mathbb{N}}$, also equipped with the product topology.

A homeomorphic space can be constructed by using the projective limit of the diagram

$$\dots \xrightarrow{p_4} \prod_{k=1}^4 [0, \frac{1}{k}] \xrightarrow{p_3} \prod_{k=1}^3 [0, \frac{1}{k}] \xrightarrow{p_2} \prod_{k=1}^2 [0, \frac{1}{k}] \xrightarrow{p_1} [0, 1],$$

where we now make “higher dimensions smaller”. The projective limit of this diagram is $\prod_{k=1}^{\infty} [0, \frac{1}{k}]$, which is known as the **Hilbert cube**.²⁶ Since there are homeomorphisms $[0, 1] \cong [0, \frac{1}{k}]$ for every positive $k \in \mathbb{N}$, the universal property of the product and the projective limits can be used to show that the projective limits of the two diagrams are also homeomorphic.

Instead of proving directly that the above projective limits are given by the corresponding infinite product, let’s do this abstractly in any category in which the relevant products exist. So suppose that we are given a sequence of objects $(A_n)_{n \in \mathbb{N}}$ in any category such that both the finite products $\prod_{i=1}^n A_i$ and the infinite product $\prod_{i=1}^{\infty} A_i$ exist. Then let

$$p_n : \prod_{i=1}^{n+1} A_i \longrightarrow \prod_{i=1}^n A_i$$

²⁶Compare with [Hilbert’s Hotel](#).

be the morphisms that forget the last factor of the product, formally defined as either the product projection of the binary product (4.6) or equivalently via the universal property of the n -fold product as unique morphism $\prod_{i=1}^{n+1} A_i \rightarrow \prod_{i=1}^n A_i$ whose i -th component is the projection onto the i -th factor for $i = 1, \dots, n$. Let the

$$u_n : \prod_{i=1}^{\infty} A_i \longrightarrow \prod_{i=1}^n A_i$$

be defined analogously. Then we consider the limit of the diagram

$$\dots \xrightarrow{p_3} \prod_{i=1}^3 A_i \xrightarrow{p_2} \prod_{i=1}^2 A_i \xrightarrow{p_1} A_1.$$

The $(u_n)_{n \in \mathbb{N}}$ form a cone over this diagram, as illustrated by

$$\begin{array}{ccccc} & & \prod_{i=1}^{\infty} A_i & & \\ & \swarrow u_3 & \downarrow u_2 & \searrow u_1 & \\ \dots & & & & \\ \dots & \xrightarrow{p_3} & \prod_{i=1}^3 A_i & \xrightarrow{p_2} & \prod_{i=1}^2 A_i & \xrightarrow{p_1} & A_1. \end{array}$$

The claim that they form a cone asserts the commutativity of the triangles. This is straightforward to prove by noting that the commutativity holds after projecting to each factor (compare Lemma 4.1.8 for the binary case).

Lemma 4.3.5. *This cone is a limit cone.*

We present the proof in somewhat abbreviated form; with a good understanding of products, you will easily be able to fill in the details yourself.

Proof sketch. Suppose that we are given any other cone consisting of morphisms (f_n) like this,

$$\begin{array}{ccccc} X & \overset{\exists! g}{\dashrightarrow} & \prod_{i=1}^{\infty} A_i & & \\ \downarrow f_3 & \searrow f_2 & \downarrow u_2 & \searrow u_1 & \\ \dots & & & & \\ \dots & \xrightarrow{p_3} & \prod_{i=1}^3 A_i & \xrightarrow{p_2} & \prod_{i=1}^2 A_i & \xrightarrow{p_1} & A_1. \end{array}$$

Then by the universal property of the infinite product $\prod_{i=1}^{\infty} A_i$, the morphism g we are looking for is uniquely determined by its components g_n . But these are already determined by the given maps to the finite products, and this shows the uniqueness of g .

Existence holds for similar reasons as follows. The composites

$$X \xrightarrow{f_k} \prod_{i=1}^k A_i \longrightarrow A_n$$

for $k \geq n$, where the second arrow is the projection, are seen to be independent of k by the assumption that the (f_n) form a cone. Using these composites as the components of g , we obtain a morphism $g : X \rightarrow \prod_{i=1}^{\infty} A_i$ that has the required properties. $\square$

We turn to an interesting example of a projective limit which occurs in probability theory.

Example 4.3.6 (Kolmogorov extension theorem). Let **Meas** be the category of measurable spaces. By this we mean the following:

- ▷ Objects are measurable spaces: these are pairs (X, Σ_X) where X is a set and $\Sigma_X \subseteq 2^X$ is a σ -algebra on X .
- ▷ Morphisms $(X, \Sigma_X) \rightarrow (Y, \Sigma_Y)$ are measurable functions $f : X \rightarrow Y$, meaning that $f^{-1}(S) \in \Sigma_X$ for every $S \in \Sigma_Y$.

We often leave the σ -algebra implicit and just write X for a measurable space. For example, $\mathbb{R}$ denotes the measurable space given by the real numbers equipped with the Borel σ -algebra.

Similar to what we did for **Top** in Example 4.1.4, one can show that products with countably many factors exist in **Meas** and are given by the Cartesian product of sets equipped with the product σ -algebra.²⁷ In particular, we can form the product $\mathbb{R}^{\mathbb{N}}$ of countably many copies of $\mathbb{R}$. A probability measure on $\mathbb{R}^{\mathbb{N}}$ corresponds to the joint distribution of countably many random variables $X_1, X_2, \dots$ taking values in $\mathbb{R}$.

What does this have to do with projective limits? To see what their significance might be, consider the following version of the **Kolmogorov extension theorem**:

The probability measures μ on $\mathbb{R}^{\mathbb{N}}$ are in bijection with the collections of probability measures $(\mu_n)_{n \in \mathbb{N}}$ on $\mathbb{R}^n$ such that for every $n \in \mathbb{N}$, the measure μ_n is the marginal of μ_{n+1} on the first n coordinates.

The similarity with the universal property of projective limits might already be clear: the (μ_n) are like a cone over a diagram consisting of the objects $\mathbb{R}^n$, where the compatibility condition on the measures corresponds to the compatibility condition on the legs of a cone. The induced measure μ corresponds to the induced dashed morphism in Figure 4.

So is this just an analogy, or does the Kolmogorov extension theorem really encode the universal property of a projective limit? To see that it really does encode a universal property, we still need to talk about probability measures in terms of the categorical structure. If we assign to every measurable space X the set

$$PX := \{\text{probability measures on } X\},$$

then $P : \mathbf{Meas} \rightarrow \mathbf{Set}$ becomes a functor if we define it on morphisms $f : X \rightarrow Y$ as taking the **pushforward** of a measure,

$$\begin{aligned} Pf : PX &\longrightarrow PY \\ \mu &\longmapsto f_*\mu, \end{aligned}$$

where $f_*\mu$ is the measure on Y defined by $(f_*\mu)(S) = \mu(f^{-1}(S))$ for every $S \in \Sigma_Y$. In case that you're comfortable with basic measure theory, you should prove that this indeed defines a functor!

In particular, we can apply the functor P to the projection maps $\mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$. This amounts precisely to the formation of the marginal distributions that appear in the statement of the extension theorem. Doing this for every n produces a diagram in **Set** given by

$$\dots \longrightarrow P(\mathbb{R}^3) \longrightarrow P(\mathbb{R}^2) \longrightarrow P(\mathbb{R}).$$

The statement of the Kolmogorov extension theorem is now that the limit of this diagram exists and is given by $P(\mathbb{R}^{\mathbb{N}})$, where the legs of the limit cone are similarly given by the pushforward maps along the projections $\mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^n$.

²⁷The details are not important for us here: our goal is to illustrate and understand the significance of projective limits, not to develop the theory of measurable spaces.

This ends our discussion of projective limits as a class of limits of a particular shape.

What can we say about the existence of limits in general? The following result on how to construct limits in **Set** not only provides important intuition, but we will see that it works in pretty much the same way in *any* category!

Proposition 4.3.7. *In **Set**, every diagram has a limit.*

Proof. We use the same notation for a given diagram as in Definition 4.3.1. Then the following considerations let us determine what the limit should be. Using a singleton set $X = 1$ in the universal property of Definition 4.3.1 shows that $\mathbf{Set}(1, L)$ must be in bijection with the set of cones $(f_v : 1 \rightarrow \ell(v))_{v \in V}$. But we can trivially identify $\mathbf{Set}(1, L)$ with L itself. Therefore we can conclude that the limit L should be given by the set of cones with apex 1!

Inspired by this idea, let us try to take

$$L := \left\{ (a_v)_{v \in V} \in \prod_{v \in V} \ell(v) \mid a_{t(e)} = \ell(e)a_{s(e)} \forall e \in E \right\}, \quad (4.15)$$

where we think of such a family $(a_v)_{v \in V}$ as a consistent choice of elements from each set in the diagram, where the consistency conditions are determined by the edges of the diagram as indicated and state that the elements a_v should map to each other upon applying any of the maps that form part of the diagram.

Let us prove that L is indeed the limit of the diagram, where the limit cone has legs $u_w : L \rightarrow \ell(w)$ given simply by projecting, meaning

$$u_w((a_v)_{v \in V}) := a_w.$$

To prove the universal property, we need to assume that we have some cone $(f_v : X \rightarrow \ell(v))_{v \in V}$ and show that it factorizes uniquely across the limit cone. But this is quite clear: the induced morphism must be

$$\begin{aligned} g : X &\longrightarrow L \\ x &\longmapsto (f_v(x))_{v \in V} \end{aligned}$$

since only this choice satisfies $u_w g = f_w$ for all $w \in V$. The remaining problem is to prove that the family of elements $(f_v(x))_{v \in V}$ indeed belongs to L , meaning that it satisfies the equations

$$f_{t(e)}(x) = \ell(e)(f_{s(e)}(v)) \quad \forall e \in E.$$

But this is precisely because of the assumption that the $(f_v)_{v \in V}$ form a cone, namely $f_{t(e)} = \ell(e)f_{s(e)}$. $\square$

The construction of L in the proofs gives a concrete description of the limit as a set that is often useful in practice. For example a projective limit, like in Figure 4, can in **Set** be constructed as the set of “compatible sequences”

$$L = \left\{ (a_1, a_2, a_3, \dots) \in \prod_{k \in \mathbb{N}} A_k \mid m_k(a_{k+1}) = a_k \forall k \in \mathbb{N} \right\}.$$

Very much the same construction can be used in many other categories like **Top** or **Grp**.

Even more generally, it is instructive to see how the proof of Proposition 4.3.7 can be “internalized” into any category with products and equalizers. This leads to the following result.

Theorem 4.3.8. *Let $\mathcal{C}$ be a category which has equalizers and a product for any family of objects. Then $\mathcal{C}$ has all limits.*

Proof. The idea is to generalize the construction of the limit (4.15) to the general abstract setting of any category $\mathcal{C}$. Looking at this construction reveals that it requires the possibility to talk about the product $\prod_{v \in V} \ell(v)$ and then impose a family of additional equations indexed by the set of edges E . Instead of considering this as a family of equations $a_{t(e)} = \ell(e)a_{s(e)}$ —namely one for each $e \in E$ —we can consider it as a *single* equation between families of elements,

$$(a_{t(e)})_{e \in E} = (\ell(e)(a_{s(e)}))_{e \in E}.$$

Like this, each side now is viewed as an element of $\prod_{e \in E} \ell(t(e))$. With this in mind, the following proof is precisely a translation of our proof in **Set**.

This indicates that we should look at an equalizer of a pair of morphisms between $\prod_{v \in V} \ell(v)$ and $\prod_{e \in E} \ell(t(e))$,

$$L \xrightarrow{i} \prod_{v \in V} \ell(v) \xrightleftharpoons[\text{move}]{\text{stay}} \prod_{e \in E} \ell(t(e)).$$

Here, the morphisms **stay** and **move** are the unique morphisms with components²⁸

$$\text{stay}_e = p_{t(e)}, \quad \text{move}_e = \ell(e)p_{s(e)},$$

where the p 's are the product projections of $\prod_{v \in V} \ell(v)$. In other words, using the universal property of the product $\prod_{e \in E} \ell(t(e))$, they are the unique morphisms that make the diagrams

$$\begin{array}{ccc} \prod_{v \in V} \ell(v) & \xrightarrow{\text{stay}} & \prod_{e \in E} \ell(t(e)) \\ & \searrow p_{t(e')} & \swarrow q_{e'} \\ & \ell(t(e')) & \end{array} \qquad \begin{array}{ccc} \prod_{v \in V} \ell(v) & \xrightarrow{\text{move}} & \prod_{e \in E} \ell(t(e)) \\ p_{s(e')} \downarrow & & \downarrow q_{e'} \\ \ell(s(e')) & \xrightarrow{\ell(e')} & \ell(t(e')) \end{array} \quad (4.16)$$

commute for all $e' \in E$, where the q 's are the product projections of $\prod_{e \in E} \ell(t(e))$.

In this way, we obtain a cone $(u_v : L \rightarrow \ell(v))_{v \in V}$ with components

$$u_v := p_v i.$$

The compatibility equations $u_{t(e)} = \ell(e)u_{s(e)}$ follow as in

$$u_{t(e)} = p_{t(e)} i = \ell(e)p_{s(e)} i = \ell(e)u_{s(e)}.$$

Here, the second equation holds because it is equivalent to $\text{stay} \circ i = \text{move} \circ i$ by the definition of **stay** and **move** and the universal property of $\prod_{e \in E} \ell(t(e))$. More formally, by (4.16) and $\text{stay} \circ i = \text{move} \circ i$, we have

$$p_{t(e)} i = q_e \circ \text{stay} \circ i = q_e \circ \text{move} \circ i = \ell(e)p_{s(e)} i.$$

and therefore the compatibility equations follow.

The proof of the universal property is also analogous to the proof we conducted for **Set** in Proposition 4.3.7, but a bit more abstract. So let $(f_v : X \rightarrow \ell(v))_{v \in V}$ be any cone. Then the tupling of its legs is the unique morphism $\hat{f} : X \rightarrow \prod_{v \in V} \ell(v)$ whose components are the f_v , meaning the unique morphism such that

$$\begin{array}{ccc} X & \xrightarrow{\hat{f}} & \prod_{v \in V} \ell(v) \\ & \searrow f_w & \swarrow p_w \\ & \ell(w) & \end{array} \quad (4.17)$$

²⁸We write stay_e for the composition of **stay** with the product projection to $A_{t(e)}$, and likewise for move_e .

commutes for all $w \in V$. We have

$$\text{stay} \circ \hat{f} = \text{move} \circ \hat{f}, \quad (4.18)$$

since looking at the component indexed by $e \in E$, which means composing with the projection q_e , we have

$$\begin{aligned} q_e \circ \text{stay} \circ \hat{f} &= p_{t(e)} \hat{f} \\ &= f_{t(e)} \\ &= \ell(e) f_{s(e)} \\ &= \ell(e) p_{s(e)} \hat{f} \\ &= q_e \circ \text{move} \circ \hat{f}, \end{aligned}$$

where we have used (4.16) and (4.17) as well as the assumption that $(f_v)_{v \in V}$ is a cone. The claimed (4.18) then follows by the uniqueness part of the universal property of the product $\prod_{e \in E} \ell(t(e))$. We therefore obtain an induced morphism $g : X \rightarrow L$ from the universal property of the equalizer, which is the unique morphism such that

$$ig = \hat{f}.$$

By again reasoning in components, it is not difficult to see that the required equations $f_v = u_v g$ are satisfied, and that g is unique with these properties. More formally, for the g constructed as above, we get

$$u_v g = p_v ig = p_v \hat{f} = f_v,$$

as required.

For the uniqueness, suppose that $g' : X \rightarrow L$ satisfies $u_v g' = f_v$ for all $v \in V$ as well. Then we have

$$p_v ig' = f_v = p_v ig \quad \forall v \in V,$$

and therefore the uniqueness part of the universal property of the product $\prod_{v \in V} \ell(v)$ implies that $ig' = ig$. We then obtain the required $g' = g$ by the fact that i is a monomorphism (Lemma 4.2.4). $\square$

4.4 Colimits

Thanks to the principle of duality (Slogan 2.5.5), the concept of limit has a dual concept called **colimit**: a colimit in $\mathcal{C}$ is defined simply as a limit in $\mathcal{C}^{\text{op}}$. One then speaks of *cocones* instead of cones. Since the general theory of colimits is just dual to that of limits, we can remain brief. The only essential difference lies in how these notions behave in those categories that are of interest in mathematical practice.

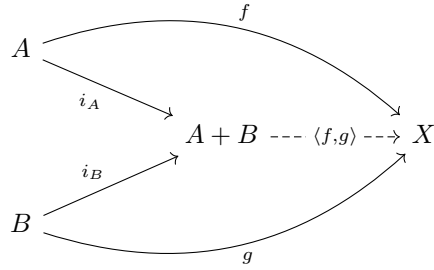
For simplicity and brevity, we limit ourselves to the special case of coproducts, which are the dual concept to products.

Definition 4.4.1. *Given a category $\mathcal{C}$ and objects $A, B \in \mathcal{C}$, their **coproduct** is an object $A + B \in \mathcal{C}$ together with morphisms*

$$i_A : A \longrightarrow A + B, \quad i_B : B \longrightarrow A + B \quad (4.19)$$

such that for every object $X \in \mathcal{C}$ and every pair of morphisms $f : A \rightarrow X$ and $g : B \rightarrow X$, there is

a unique morphism $\langle f, g \rangle : A + B \rightarrow X$ such that the diagram



commutes. We call the morphisms i_A and i_B the **coprojections** and the morphism $\langle f, g \rangle$ the **copairing** of f and g .

We (co)limit ourselves to the example of **Set**.

Example 4.4.2. Coproducts in **Set** exist and are given by taking $A + B$ to be the *disjoint union* of A and B . The coprojections (4.19) are the canonical inclusions of A and B into their disjoint union.

It is quite clear that the universal property holds: any two functions $A \rightarrow X$ and $B \rightarrow X$ can be uniquely “glued” to a function $A + B \rightarrow X$. In particular, a function $A + B \rightarrow X$ is uniquely determined by what it does on A and B separately. These are basic principles for working with disjoint unions that we frequently use in mathematical practice.

5 Adjoint functors

A common situation in mathematics is that one studies two categories such that every object in one has a “counterpart” in the other and vice versa. This is the topic of adjoint functors and adjunctions.

5.1 Definition and examples

To keep some concrete example in mind, let us consider the forgetful functor $U : \mathbf{Vect}_k \rightarrow \mathbf{Set}$, so that the “counterpart in $\mathbf{Set}$ ” of a vector space is its underlying set. In the converse direction, what is the “vector space counterpart” of a set A ? Well, the only canonical and interesting vector space that we can build out of A is the *free vector space generated by A* , by which we mean

$$k[A] := \{v : A \rightarrow k \text{ of finite support}\}, \quad (5.1)$$

where finite support means that $v(a) \neq 0$ for only finitely many a , and $k[A]$ becomes a vector space with respect to the usual pointwise addition and scalar multiplication of functions,

$$\begin{aligned} (v + w)(a) &:= v(a) + w(a), \\ (\lambda v)(a) &:= \lambda v(a). \end{aligned}$$

The right way to think about $k[A]$ is that it is a canonical vector space which has the set A as a basis: an element v can be thought of as a map which assigns to every basis element the coefficient of the vector in the basis. This is consistent with the following fact: if for any $a \in A$ we write $\delta_a : A \rightarrow k$ for the function which is 1 on a itself and 0 on all other elements, then every element $v \in k[A]$ can be written uniquely as

$$v = \sum_{a \in A} v(a) \delta_a, \quad (5.2)$$

where the sum only has finitely many nonzero terms by the finite support assumption. Thus the set $\{\delta_a : a \in A\} \cong A$ is a basis of $k[A]$, and because of this we can think of $k[A]$ as the vector space spanned by the elements of A . Often $k[A]$ is called the vector space of “formal linear combinations” of the elements of A . We prefer the definition (5.1) because it is more precise: making the concept of “formal linear combination” rigorous is exactly what (5.1) achieves. Based on (5.2), we can use the more intuitive formal linear combination notation from now on.

Furthermore, when we have a map $f : A \rightarrow B$, then we obtain an induced linear map

$$\begin{aligned} k[f] : k[A] &\longrightarrow k[B] \\ \sum_{a \in A} v(a) \delta_a &\longmapsto \sum_{a \in A} v(a) \delta_{f(a)} = \sum_{b \in B} \left(\sum_{a \in f^{-1}(b)} v(a) \right) \delta_b. \end{aligned}$$

In other words, we simply apply f to the basis elements δ_a , and consider the unique linear extension from there. In this way, we obtain a functor

$$k[-] : \mathbf{Set} \longrightarrow \mathbf{Vect}_k,$$

and for any $A \in \mathbf{Set}$, we think of $k[A]$ as the “vector space counterpart” of A .

Anyway, how do our two concepts of “counterpart” relate to each other? The relevant observation here is that maps $A \rightarrow U(V)$ correspond to linear maps $k[A] \rightarrow V$, meaning that there is a bijection

$$\mathbf{Set}(A, U(V)) \cong \mathbf{Vect}_k(k[A], V). \quad (5.3)$$

Indeed from right to left, we just restrict a linear map $f : k[A] \rightarrow V$ to the basis elements δ_a for $a \in A$, and this defines a map $A \rightarrow U(V)$,

$$\begin{aligned} \text{Vect}_k(k[A], V) &\longrightarrow \text{Set}(A, U(V)) \\ f &\longmapsto (a \mapsto f(\delta_a)). \end{aligned}$$

In the other direction, we can extend a map $A \rightarrow U(V)$ to a linear map $k[A] \rightarrow V$ by sending δ_a to the value of the map on a , and then using the fact that a map on a basis uniquely extends to a linear map on the whole vector space. Explicitly, this is given by

$$\begin{aligned} \text{Set}(A, U(V)) &\longrightarrow \text{Vect}_k(k[A], V) \\ g &\longmapsto \left(\sum_{a \in A} v(a) \delta_a \mapsto \sum_{a \in A} v(a) g(a) \right). \end{aligned}$$

It is not difficult to see that these two constructions are inverses of each other, so we have established the bijection.

This suggests the following general definition.

Definition 5.1.1. For functors $F : \mathcal{C} \rightarrow \mathcal{D}$ and $G : \mathcal{D} \rightarrow \mathcal{C}$, an **adjunction** is a family of bijections

$$\mathcal{C}(X, GY) \cong \mathcal{D}(FX, Y) \quad (5.4)$$

natural in $X \in \mathcal{C}$ and $Y \in \mathcal{D}$.

Here, the naturality in X means that for every $h \in \mathcal{C}(X, X')$, the resulting diagram

$$\begin{array}{ccc} \mathcal{C}(X', GY) & \xrightarrow{\cong} & \mathcal{D}(FX', Y) \\ \downarrow - \circ h & & \downarrow - \circ Fh \\ \mathcal{C}(X, GY) & \xrightarrow{\cong} & \mathcal{D}(FX, Y) \end{array} \quad (5.5)$$

commutes. Similarly, the naturality in Y means that for every $k \in \mathcal{D}(Y, Y')$, the diagram

$$\begin{array}{ccc} \mathcal{C}(X, GY) & \xrightarrow{\cong} & \mathcal{D}(FX, Y) \\ Gk \circ - \downarrow & & \downarrow k \circ - \\ \mathcal{C}(X, GY') & \xrightarrow{\cong} & \mathcal{D}(FX, Y') \end{array} \quad (5.6)$$

commutes. You may want to convince yourself that these diagrams indeed commute for the bijections (5.3) constructed above.

Example 5.1.2. The forgetful functor $U : \mathbf{Top} \rightarrow \mathbf{Set}$ is part of two different adjunctions:

- (a) Let $\text{Disc} : \mathbf{Set} \rightarrow \mathbf{Top}$ be the functor which equips every set A with the discrete topology, that is the topology where every subset is open. Then for every topological space X , every map $\text{Disc}(A) \rightarrow X$ is continuous. This means that we have a bijection

$$\text{Set}(A, U(X)) \cong \text{Top}(\text{Disc}(A), X) \quad (5.7)$$

which simply takes every map to itself. Since the naturality diagrams (5.5) and (5.6) hold trivially, we are dealing with an adjunction.

- (b) Let $\text{Codisc} : \mathbf{Set} \rightarrow \mathbf{Top}$ be the functor which equips every set A with the indiscrete topology, that is the topology in which only the empty set and A itself are open. Then for every topological space X , every map $X \rightarrow \text{Codisc}(A)$ is continuous. Just as in the previous item, this means that we have an adjunction with components

$$\text{Top}(X, \text{Codisc}(A)) \cong \text{Set}(U(X), A). \quad (5.8)$$

Example 5.1.3 (Galois connections). Let $(X, \leq)$ and $(Y, \leq)$ be preordered sets, considered as categories as described in Example 2.1.7. Then an adjunction between X and Y consists of monotone maps

$$f : X \rightarrow Y, \quad g : Y \rightarrow X$$

such that for all $x \in X$ and $y \in Y$, we have

$$x \leq g(y) \iff f(x) \leq y. \quad (5.9)$$

Indeed since the hom-sets are all either singletons or empty, the bijections (5.4) merely amount to the condition that if the left-hand side or right-hand side contains an element, then so does the other side, and the naturality conditions (5.5) and (5.6) hold automatically. This translates exactly into our inequalities (5.9). Such a situation is called a **Galois connection**. They arise in many different ways throughout mathematics.

- (a) Let $h : A \rightarrow B$ be any map between sets A and B . Then the operations of forming the image and preimage of a subset along h form a Galois connection between the power sets,

$$\begin{array}{ccc} & S \mapsto f(S) & \\ 2^A & \xrightarrow{\quad} & 2^B \\ & T \mapsto f^{-1}(T) & \end{array}$$

due to the well-known equivalence

$$S \subseteq f^{-1}(T) \iff f(S) \subseteq T.$$

- (b) Let $\lfloor \cdot \rfloor : \mathbb{R} \rightarrow \mathbb{Z}$ the map which assigns to every $r \in \mathbb{R}$ its integer part $\lfloor r \rfloor \in \mathbb{Z}$. Then together with the inclusion map $i : \mathbb{Z} \rightarrow \mathbb{R}$, we obtain an adjunction because

$$n \leq \lfloor r \rfloor \iff i(n) \leq r$$

for all $n \in \mathbb{Z}$ and $r \in \mathbb{R}$.

- (c) Let $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ and let X be any finite or infinite set. In optimization, one is interested in computing the map

$$\sup : \mathbf{Set}(X, \overline{\mathbb{R}}) \longrightarrow \overline{\mathbb{R}}$$

which assigns to every function $X \rightarrow \overline{\mathbb{R}}$ the supremum of its values. In the other direction, we have a rather trivial monotone map assigning to every number the corresponding constant function,

$$\begin{aligned} \text{const} : \overline{\mathbb{R}} &\longrightarrow \mathbf{Set}(X, \overline{\mathbb{R}}) \\ r &\longmapsto (x \mapsto r). \end{aligned}$$

Then for every $f \in \mathbf{Set}(X, \overline{\mathbb{R}})$ and $r \in \overline{\mathbb{R}}$, the definition of supremum shows that

$$f \leq \text{const}(r) \iff \sup(f) \leq r,$$

since this equivalence asserts precisely that $\sup(f)$ is the smallest upper bound of all the values of f . We are therefore dealing with a Galois connection. If we use the infimum instead of the supremum, then we similarly get

$$r \leq \inf(f) \iff \text{const}(r) \leq f.$$

- (d) As the name indicates, a Galois connection also features prominently in Galois theory. In fact, the fundamental theorem of Galois theory *is* a Galois connection in the following sense: for any field k , we have a Galois connection

$$\begin{array}{ccc} & \text{Stab} & \\ \text{subfields of } k & \xrightarrow{\quad} & \text{subgroups of } \text{Aut}(k)^{\text{op}} \\ & \xleftarrow{\quad} & \\ & \text{Fix} & \end{array} \quad (5.10)$$

where **Stab** sends a subfield $k' \subseteq k$ to those automorphisms $\varphi : k \rightarrow k$ which stabilize it pointwise, and **Fix** sends a subgroup $G \subseteq \text{Aut}(k)$ to the subfield of elements that are fixed by G . Then it is quite obvious that

$$k' \subseteq \text{Fix}(G) \iff G \subseteq \text{Stab}(k').$$

because both sides are equivalent to the condition that $\varphi(x) = x$ for all $x \in k'$ and $\varphi \in G$. In light of the “op” in (5.10), this indeed defines a Galois connection. This adjunction is the “easy” part of Galois theory, in the sense that it holds essentially by definition of **Stab** and **Fix** and does not require any knowledge of algebra. Thanks to this simplicity, analogous Galois connections exist and are useful in many other contexts.

Remark 5.1.4. When there exists an adjunction involving functors F and G , then we sometimes denote this by

$$F \dashv G.$$

In practice, this is usually sufficient to specify an adjunction: there typically is only one natural (in both the categorical and the common language sense of the word) choice for the bijections (5.4) once the categories and functors are known. Given such an adjunction, we call F the **left adjoint** of G , and likewise G the **right adjoint** of F . One also often uses these terms as adjectives, saying that F is *left adjoint to* G , and likewise that G is *right adjoint to* F .

For example, we can abbreviate Example 5.1.2 as

$$\text{Disc} \dashv U \dashv \text{Codisc},$$

and we can say that **Disc** is the left adjoint of U , which in turn is the left adjoint of **Codisc**.

We have ended the class here

5.2 Adjunctions via universal properties

A glance at (5.4) shows that for every $Y \in \mathcal{D}$, the object GY represents the functor

$$\mathcal{D}(F-, Y) : \mathcal{C}^{\text{op}} \rightarrow \text{Set}.$$

As is generally the case with representable functors (see Remark 3.3.2), the universal element is the element of the functor which corresponds to the identity on the universal object. In the present case, this is a morphism

$$\varepsilon_Y : FGY \longrightarrow Y.$$

This morphism is called the **counit** of the adjunction, or more precisely the *component* of the counit at $Y \in \mathcal{D}$ because of the following:

Lemma 5.2.1. *The counit is a natural transformation $\varepsilon : FG \Rightarrow \text{id}_{\mathcal{D}}$.*

Proof. When we consider any $\ell : Y \rightarrow Y'$ in $\mathcal{D}$, then we need to show that the diagram

$$\begin{array}{ccc} FGY & \xrightarrow{FG\ell} & FGY' \\ \varepsilon_Y \downarrow & & \downarrow \varepsilon_{Y'} \\ Y & \xrightarrow{\ell} & Y' \end{array}$$

commutes. But this follows from the commutativity of (5.5) and (5.6), which for suitable choices of the objects and morphisms involved give the diagram

$$\begin{array}{ccc} \mathcal{C}(GY, GY) & \xrightarrow{\cong} & \mathcal{D}(FGY, Y) \\ G\ell \circ - \downarrow & & \downarrow \ell \circ - \\ \mathcal{C}(GY, GY') & \xrightarrow{\cong} & \mathcal{D}(FGY, Y') \\ \uparrow - \circ G\ell & & - \circ FG\ell \uparrow \\ \mathcal{C}(GY', GY') & \xrightarrow{\cong} & \mathcal{D}(FGY', Y') \end{array}$$

Now starting with id_{GY} top left and with $\text{id}_{GY'}$ bottom left both produce Gh at centre left. Therefore both initial conditions yield the same element centre right, and this amounts to the equation

$$\ell \circ \varepsilon_Y = \varepsilon_{Y'} \circ FG\ell,$$

as was to be shown. □

Similarly, an adjunction has a **unit**

$$\eta : \text{id}_{\mathcal{C}} \Rightarrow GF$$

whose component at an object $X \in \mathcal{C}$ is the counterpart $\eta_X : X \rightarrow GFX$ of id_{FX} under (5.4). Let us check out what the unit and counit are in our examples in order to gain some intuition for them.

Example 5.2.2. In the case of the adjunction between sets and vector spaces given by (5.3), the counit at a vector space V is a linear map

$$\varepsilon_V : k[U(V)] \longrightarrow V.$$

Which linear map is it exactly? Well, since ε_V is the counterpart of id_V under (5.3), it is the unique map which maps every basis vector to itself, and consequently maps a formal linear combination to the *actual* linear combination that it denotes,

$$\varepsilon_V \left(\sum_{x \in V} v(x) \delta_x \right) = \sum_{x \in V} v(x) x. \quad (5.11)$$

The unit at a set A is easier: it is the map

$$\eta_A : A \longrightarrow U(k[A])$$

which sends every element $a \in A$ to the corresponding basis vector $\delta_a \in k[A]$. It should be clear that both the unit and the counit encode important aspects of the vector space structure.

Example 5.2.3. The unit and counit for the adjunction $\text{Disc} \dashv U$ from (5.7) are both given by the identity on points. With $U(\text{Disc}(A)) = A$ in mind, the unit components are $\eta_A = \text{id}_A$. The counit components are the maps $\varepsilon_X : \text{Disc}(U(X)) \rightarrow X$ which are also the identity on points, but where the topology changes from the discrete topology to the given topology on X .

Similarly, for the adjunction $U \dashv \text{Codisc}$ from (5.8), the counit components are

$$\varepsilon_A = \text{id}_A : U(\text{Codisc}(A)) \rightarrow A,$$

while the unit components are

Example 5.2.4. For a Galois connection as in Example 5.1.3, the counit witnesses the inequalities

$$f(g(y)) \leq y \quad \forall y \in Y,$$

and the unit similarly witnesses the inequalities

$$x \leq g(f(x)) \quad \forall x \in X.$$

Proposition 5.2.5. (a) *If a functor $F : \mathcal{C} \rightarrow \mathcal{D}$ has a right adjoint, then this right adjoint is unique up to natural isomorphism.*

(b) *If a functor $G : \mathcal{C} \rightarrow \mathcal{D}$ has a left adjoint, then this left adjoint is unique up to natural isomorphism.*

Proof. The two statements are duals of each other, so we only prove the first.

Thus if $G' : \mathcal{D} \rightarrow \mathcal{C}$ is another right adjoint to F , then by the uniqueness of the representing object from $\square$

Many examples of adjunctions arise from the following situation.

Definition 5.2.6. *Given a category $\mathcal{C}$, a **full subcategory** $\mathcal{D}$ consists of a subcollection of objects $\text{Ob}(\mathcal{D}) \subseteq \text{Ob}(\mathcal{C})$ such that $\mathcal{D}$ contains all morphisms between these objects, that is*

$$\mathcal{D}(X, Y) = \mathcal{C}(X, Y) \quad \forall X, Y \in \text{Ob}(\mathcal{D}).$$

In mathematical practice, one often encounters situations where one wants to “force” an object to belong to a certain full subcategory. Of course this requires modifying the object, in ways which typically either amount to some kind of completion or to forming a quotient which imposes additional equations. Let us give an example of each kind.

Example 5.2.7 (Completion of a metric space). For a metric space X with metric d_X , recall that the **completion** is a complete metric space $\bar{X}$ into which X embeds as a dense subspace and whose elements are Cauchy sequences.

To see how to treat the completion in categorical terms, let Met_1 be the category of metric spaces and **contractive maps**, where a map $f : X \rightarrow Y$ is contractive if

$$d_Y(f(x_1), f(x_2)) \leq d_X(x_1, x_2) \quad \forall x_1, x_2 \in X.$$

Let furthermore $\text{cMet}_1 \subseteq \text{Met}_1$ be the full subcategory of complete metric spaces. Then for every $X \in \text{Met}_1$ and $Y \in \text{cMet}_1$, we have a bijection

$$\text{Met}_1(X, Y) \cong \text{cMet}_1(\bar{X}, Y),$$

because every contractive $f : X \rightarrow Y$ can be extended uniquely to $\bar{f} : \bar{X} \rightarrow Y$ by defining $\bar{f}((x_n)_{n \in \mathbb{N}})$ for a Cauchy sequence $(x_n)_{n \in \mathbb{N}}$ to be the limit of the Cauchy sequence $(f(x_n))_{n \in \mathbb{N}}$ in Y . It is easy to see that this bijection is natural in X and Y as required, and therefore the inclusion functor $\text{cMet}_1 \hookrightarrow \text{Met}_1$

Example 5.2.8 (Abelianization of a group). Let

6 Miscellaneous topics

We probably will not get to this

6.1 Equivalence of categories

6.2 The Yoneda embedding

We make a small detour by deriving another consequence of the Yoneda lemma. First, it may help to recall *Cayley's theorem* from group theory. Let us state and prove the

Proposition 6.2.1 (Cayley's theorem). *Let G be a group. Then the map*

$$\begin{aligned}\phi : G &\longrightarrow \text{Set}(G, G) \\ g &\longmapsto (x \mapsto gx)\end{aligned}$$

is an injective monoid homomorphism and embeds G into the group of permutations of the set G .

Theorem 6.2.2 (Yoneda embedding). *Let $\mathcal{C}$ be a locally small category. Then the functor*

$$\begin{aligned}\mathcal{C} &\longrightarrow [\mathcal{C}, \text{Set}] \\ X &\longmapsto \mathcal{C}(X, -)\end{aligned}$$

is full and faithful.

Index of Categories

BG , 7

Bord_n , 12

Cat , 18

cMet1cMet_1 , 62

Euc , 17

Fld , 39

Functor category, $[\mathcal{C}, \mathcal{D}]$, 25

Grp , 6

Hilb , 14

Mat_K , 6

Meas , 52

Met_1 , 62

Met_{cont} , 12

Mon , 8

Opposite category, $\mathcal{C}^{\text{op}}$, 13

PSuccSet , 34

Rel , 13

Set , 5

Top , 6

Vect_K , 15

References

- [1] Steve Awodey. *Category Theory*. Oxford University Press, 2010.
- [2] Martin Brandenburg. *Einführung in die Kategorientheorie. Mit ausführlichen Erklärungen und zahlreichen Beispielen*. Springer, 2017.
- [3] Tobias Fritz, Paolo Perrone, and Sharwin Rezagholi. “Probability, valuations, hyperspace: Three monads on Top and the support as a morphism”. In: *Mathematical Structures in Computer Science* 31.8 (2021). arxiv.org/abs/1910.03752, pp. 850–897.
- [4] Tom Leinster. *Basic Category Theory*. arxiv.org/abs/1612.09375. Cambridge University Press, 2014.
- [5] Saunders Mac Lane. *Categories for the Working Mathematician*. Springer, 1978.
- [6] Paolo Perrone. *Starting Category Theory*. The published version of this book essentially coincides with the preprint arxiv.org/abs/1912.10642. World Scientific, 2024.
- [7] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. homotopytypetheory.org/book. Institute for Advanced Study, 2013.
- [8] Emily Riehl. *Category Theory in Context*. math.jhu.edu/~eriehl/context.pdf. Dover Publications, 2016.