1 Basic category theory

1.1 Categories

1.1.1 What is a category?

It is surprisingly hard to define mathematics, but people tend to agree that mathematics has something to do with studying mathematical structures.[1] Mathematicians study lots of different mathematical structures, each of which is good for different things. Groups are good for talking about symmetries, for instance, and topological spaces are good for talking about geometrical objects.

Categories are just another type of mathematical structure, but they have a different, more abstract feel to them than groups or topological spaces. One reason for this is that they are inherently meta-mathematical: they are good for talking about other mathematical objects and the relationships between them.

For this reason, category theory has developed a reputation for being abstract and difficult to learn, which is mostly undeserved. Categories are nothing special; they are mathematical objects, just like groups or topological spaces. They are simply good for different things.

Unfortunately, there are many different conventions regarding how to typeset category theory. I will do my best to stick to the conventions used at the nLab ([?]).

Definition 1 (category)**.** A category C consists of the following pieces of data.

- A class[2] $\mathrm{Obj}(\mathsf{C})$ of *objects*.
- For every two objects $A, B \in \mathrm{Obj}(\mathsf{C})$, a class $\mathrm{Hom}(A, B)$ of *morphisms*.
- For every objects X, Y, Z, a map

  $$\mathrm{Hom}(Y, Z) \times \mathrm{Hom}(X, Y) \to \mathrm{Hom}(X, Z); \qquad (f, g) \mapsto f \circ g$$

  called the *composition* of f and g, which satisfies the following properties.

  1. This composition is associative: $(f \circ g) \circ h = f \circ (g \circ h)$.
  2. For every $A \in \mathrm{Obj}(\mathsf{C})$, there is at least one morphism id_A, called the *identity morphism* which functions as both a left and right identity with respect to the composition of morphisms, i.e.

  $$f \circ \mathrm{id}_A = f \qquad \text{amd} \qquad \mathrm{id}_A \circ g = g.$$

---

[1] Whatever they are.

[2] Note that if we choose to work within ZFCU (see appendix REF), one gets away with saying a *set* of objects. However, this doesn't really make much of a difference to the development of the theory. We will mostly avoid issues of size.

If ever it is potentially unclear which category we are talking about, we will add a subscript to Hom, writing for example $\mathrm{Hom}_{\mathsf{C}}(A, B)$ instead of $\mathrm{Hom}(A, B)$.

Notation 2. Following Aluffi ([?]), we will use a sans serif font to denote categories. For example C, Set.

One often thinks of the objects in a category as 'generalized sets,' and of morphisms as 'generalized functions.' It is therefore often useful to use functional notation $f\colon A \to B$ to describe morphisms. For example, the following two notations are equivalent:

$$f \in \mathrm{Hom}(A, B) \iff f\colon A \to B.$$

Here is something to be aware of: the identity morphism $\mathrm{id}_A\colon A \to A$ is often simply denoted by A. This is actually a good notation, and we will use it freely in later chapters. However, we will avoid it in earlier chapters since it is potentially confusing whether we are talking about A the object or A the morphism.

1.1.2 Some examples of categories

The idea of a category is very abstract, but not abstruse. The easiest way to convince oneself of this is to see some examples. The most common type of category is given by 'mathematical structures and structure-preserving maps between them,' so many of our examples will be in this vein.

Example 3 (The category of sets)**.** The prototypical category is Set, the category whose objects are sets and whose morphisms are functions between them.

In order to check that Set is indeed a category, we need to check that the axioms are satisfied. In fact, everything is more or less clear by definition.

Our category Set first needs a class of objects. Although there is no set of all sets, there certainly is a class of small sets. Therefore, we define

$$\mathrm{Obj}(\mathsf{Set}) = \text{all sets}.$$

The next thing Set needs is, for each pair of objects (i.e. sets) A and B, a class of morphisms $\mathrm{Hom}_{\mathsf{Set}}(A, B)$ between them. The morphisms between the objects in Set will consist of all set-functions between. That is,

$$\mathrm{Hom}_{\mathsf{Set}}(A, B) = \{f\colon A \to B \mid f \text{ is a function}\}.$$

Next one has to check that functions satisfy the necessary axioms. Given a function $f\colon A \to B$ and a function $g\colon B \to C$, we can compose them to get a function $g \circ f\colon A \to C$. Thus, functions can indeed be composed in the necessary way.

The composition of morphisms is associative, since composition of functions is.

Furthermore, every set A has an identity function

$$\mathrm{id}_A\colon A \to A; \qquad a \mapsto a.$$

which maps every element to itself.

Thus, Set is a category! That's all there is to it!

Example 4 (category with one object)**.** There is an important category called 1, which has one object and one morphism.

- The set of objects Obj(1) is the singleton $\{*\}$.
- The only morphism is the identity morphism $\mathrm{id}_* \colon * \to *$.

Here is a picture of the category 1.

$$\begin{tikzcd} * \arrow[loop right, "\mathrm{id}_*"] \end{tikzcd}$$

Example 5 (poset category)**.** Let $(P, \leq)$ be a poset. One can define from this a category $\mathcal{P}(P)$, whose objects are $\mathrm{Obj}(\mathcal{P}) = P$ and whose morphisms are

$$\mathrm{Hom}_{\mathcal{P}}(A, B) = \begin{cases} \{*\}, & A \leq B \\ \emptyset, & \text{otherwise.} \end{cases}$$

This immediately furnishes us with many examples of categories. For example:

- Any set
$$[n] = \{0, 1, \dots, n\}$$
is a poset with $\leq$ inherited from the integers. We can draw these categories as follows (omitting identity arrows)

$$[0]\colon\ \begin{tikzcd} 0 \end{tikzcd}\ , \qquad [1]\colon\ \begin{tikzcd} 0 \arrow[r] & 1 \end{tikzcd}\ , \qquad [2]\colon\ \begin{tikzcd} & 1 \arrow[dr] & \\ 0 \arrow[ur] \arrow[rr] & & 2 \end{tikzcd}\ , \qquad [3]\colon\ \begin{tikzcd} 0 \arrow[r] \arrow[d] \arrow[dr] & 1 \arrow[d] \arrow[dl] \\ 2 \arrow[r] & 3 \end{tikzcd}$$

- Any power set is a poset. The poset category of the power set of $\{0, 1\}$, denoted $\mathcal{P}(\{0, 1\})$ by mild abuse of notation, can be drawn as follows.

$$\begin{tikzcd} \emptyset \arrow[r] \arrow[d] \arrow[dr] & \{0\} \arrow[d] \\ \{1\} \arrow[r] & \{0, 1\} \end{tikzcd}$$

Categories capture the essence of 'things and composable maps between them.' There are many mathematical structures which naturally come with a notion of map between them, usually called homomorphisms. These structure preserving maps are almost always composable, which means it is often natural to imagine such objects as living in their own categories.

Example 6. If you know what the following mathematical objects are, it is nearly effortless to check that they are categories.

- Grp, whose objects are groups and whose morphisms are group homomorphisms.
- Ab, whose objects are abelian groups and whose morphisms are group homomorphisms.
- Ring, whose objects are rings and whose morphisms are ring homomorphisms.
- R-Mod, whose objects are modules over a ring R and whose morphisms are module homomorphisms.

- Vect_k, whose objects are vector spaces over a field k and whose morphisms are linear maps.
- $\mathsf{FinVect}_k$, whose objects are finite-dimensional vector spaces over a field k and whose morphisms are linear maps.
- k-Alg, whose objects are algebras over a field k and whose morphisms are algebra homomorphisms.

Example 7. In addition to algebraic structures, categories help to talk about geometrical structures. The following are also categories.

- Top, whose objects are topological spaces and whose morphisms are continuous maps.
- Met, whose objects are metric spaces and whose morphisms are metric maps.
- Man^p, whose objects are manifolds of class C^p and whose morphisms are p-times differentiable functions.
- $\mathsf{SmoothMfd}$, whose objects are C^∞ manifolds and whose morphisms are smooth functions.

Example 8. Here is a slightly whimsical example of a category, which is different in nature to the other categories we've looked at before. This category is important when studying group representations.

Let G be a group. We are going to create a category BG which behaves like this group.

- Our category BG has only one object, called $*$.
- The set $\mathrm{Hom}_{BG}(*, *)$ is equal to the underlying set of the group G, and for $f, g \in \mathrm{Hom}_{BG}(*, *)$, the compositions $f \circ g = f \cdot g$, where $\cdot$ is the group operation in G. The identity $e \in G$ is the identity morphism id_* on $*$.

$$\begin{tikzcd} * \arrow[loop above, "g"] \arrow[loop left, "f"] \arrow[loop right, "e=\mathrm{id}_*"] \arrow[loop below, "g\circ f"] \end{tikzcd}$$

Example 9. Let X be a topological space. We can define a category $\pi_{\leq q}(X)$ whose objects are the points of X, and whose morphisms $x \longrightarrow y$ are homotopy classes of paths $x \longrightarrow y$.

Commutative diagrams

It is often helpful to visualize objects and morphisms in categories as as graphs[3] with the objects as vertices and the morphisms as edges. This makes it possible to reason about categories using graph-like diagrams. Scroll to almost any later page in this document, and you

---

[3] Strictly speaking, they are *multidigraphs,* i.e. directed graphs which are allowed to have more than one edge between any given pair of vertices.

will likely see a diagram or two which resembles this one (copied directly from Definition ??).

$$\begin{tikzcd}
V \arrow[dr, dashed, "\exists! h"] \arrow[drr, bend left, "t"] \arrow[ddr, bend right, "s"'] & & \\
& U \arrow[r, "q"] \arrow[d, "p"'] & B \arrow[d, "g"] \\
& A \arrow[r, "f"'] & C
\end{tikzcd}$$

Category theorists frequently employ these diagrams, called *commutative diagrams,*[4] to make the ideas beind a proof more clear. Many results in category theory are most easily proved by drawing such a diagram, and then pointing at edges and vertices while repeating the phrase 'and then this goes here' and furrowing one's brow. This sort of proof is called a 'diagram chase,' and category theorists are very fond of them.

1.1.3 Building new categories from existing ones

One of the first things one learns about when talking about a mathematical structure is how to use existing ones to create new ones. For example, one can take the product of two groups to create a third, or take the quotient a vector space by a subspace. Categories are no exception. In this section, we will learn how to create new categories out of old categories.

The opposite category

The first way of doing this we'll examine creates out of any category its *opposite category,* where all the morphisms go the other way.

Definition 10 (opposite category)**.** Let C be a category. Its opposite category C^{op} is the category defined in the following way.

- The objects of C^{op} are the same as the objects of C:

$$\mathrm{Obj}(\mathsf{C}^{\mathrm{op}}) = \mathrm{Obj}(\mathsf{C}).$$

- The morphisms $A \longrightarrow B$ in C^{op} are defined to be the morphisms $B \longrightarrow A$ in C:

$$\mathrm{Hom}_{\mathsf{C}^{\mathrm{op}}}(A, B) = \mathrm{Hom}_{\mathsf{C}}(B, A).$$

That is to say, the opposite category is the category one gets by formally reversing all the arrows in a category. If $f \in \mathrm{Hom}_{\mathsf{C}}(A, B)$, i.e. $f \colon A \longrightarrow B$, then in C^{op}, $f \colon B \longrightarrow A$.

This may seem like an uninterersting definition, but having it around will make life a lot easier when it comes to defining functors (Section 1.2).

---

[4] We will define later when a diagram does or does not commute. Annoyingly, even diagrams which don't commute are often called commutative diagrams.

Product categories

One can take the Cartesian product of two sets by creating ordered pairs. One can also take the product of two categories, which works in basically the same way.

Definition 11 (product category)**.** Let C and D be categories. The product category $\mathsf{C} \times \mathsf{D}$ is the following category.

- The objects $\mathrm{Obj}(\mathsf{C} \times \mathsf{D})$ consist of all ordered pairs (C, D), where $C \in \mathrm{Obj}(\mathsf{C})$ and $D \in \mathrm{Obj}(\mathsf{D})$.
- For any two objects (C, D), and $(C', D') \in \mathrm{Obj}(\mathsf{C}\times\mathsf{D})$, the morphisms $\mathrm{Hom}_{C\times D}((C, D), (C', D'))$ are ordered pairs (f, g), where $f \in \mathrm{Hom}_{\mathsf{C}}(C, C')$ and $g \in \mathrm{Hom}_{\mathsf{D}}(D, D')$.
- Composition is taken componentwise, so that

$$(f_1, g_1) \circ (f_2, g_2) = (f_1 \circ g_1, f_2 \circ g_2).$$

- The identity morphisms are given in the obvious way:

$$\mathrm{id}_{(C,D)} = (\mathrm{id}_C, \mathrm{id}_D).$$

Product categories are best pictured in the following way. The object $(C, D) \in \mathrm{Obj}(\mathsf{C}, \mathsf{D})$ is two objects right next to each other.

$$\begin{tikzcd} C \\ D \end{tikzcd}$$

A morphism $(f, g)\colon (C, D) \longrightarrow (C', D')$ is two morphisms next to each other:

$$\begin{tikzcd} C \arrow[r, "f"] & C' \\ D \arrow[r, "g"] & D' \end{tikzcd}$$

Subcategories

Many mathematical objects have a notion of 'sub-object.' For example, groups can have subgroups, and vector spaces can have vector subspaces. Categories also have this notion. As with groups and vector spaces, the definition of a 'subcategory' is more or less obvious, but it is helpful to know exactly what one has to check to show that a subcollection of objects and morphisms from a category forms a subcategory.

Definition 12 (subcategory)**.** Let C be a category. A category S is a subcategory of C if the following conditions hold.

- The objects $\mathrm{Obj}(\mathsf{S})$ of S are a subcollection of the objects of C.
- For S, $T \in \mathrm{Obj}(\mathsf{S})$, the morphisms $\mathrm{Hom}_{\mathsf{S}}(S, T)$ are a subcollection of the morphisms $\mathrm{Hom}_{\mathsf{C}}(S, T)$ which satisfy the following.
  - For every $S \in \mathrm{Obj}(S)$, the identity $\mathrm{id}_S \in \mathrm{Hom}_{\mathsf{S}}(S, S)$.
  - For all $f \in \mathrm{Hom}_{\mathsf{S}}(S, T)$ and $g \in \mathrm{Hom}_{\mathsf{S}}(T, U)$, the composite $g \circ f \in \mathrm{Hom}_{\mathsf{S}}(S, U)$.

If S is a subcategory of C, we will write $\mathsf{S} \subseteq \mathsf{C}$.

One particularly important kind of subcategory occurs when one takes a subset of all objects in a subcategory, but keeps all morphisms between them. This is known as a *full subcategory.*

Definition 13 (full subcategory)**.** Let C be a category, $\mathrm{S} \subseteq \mathrm{C}$ a subcategory. We say that S is full in C if for every $S, T \in \mathrm{Obj}(\mathrm{S})$, $\mathrm{Hom}_{\mathrm{S}}(S,T) = \mathrm{Hom}_{\mathrm{C}}(S,T)$. That is, if there are no morphisms in $\mathrm{Hom}_{\mathrm{C}}(S,T)$ that are not in $\mathrm{Hom}_{\mathrm{S}}(S,T)$.

Example 14. Recall that Vect_k is the category of vector spaces over a field k, and $\mathsf{FinVect}_k$ is the category of finite dimensional vector spaces.

It is not difficult to see that $\mathsf{FinVect}_k \subset \mathrm{Vect}_k$: all finite dimensional vector spaces are vector spaces, and all linear maps between finite-dimensional vector spaces are maps between vector spaces. In fact, since for V and W finite-dimensional, one does not gain any maps by moving from $\mathrm{Hom}_{\mathsf{FinVect}_k}(V,W)$ to $\mathrm{Hom}_{\mathsf{Vect}_k}(V,W)$, $\mathsf{FinVect}_k$ is even a *full* subcategory of Vect_k.

1.1.4 Properties of morphisms

Category theory has many essences, one of which is as a major generalization of set theory. It is often possible to upgrade statements about functions between sets to statements about morphisms between objects in an arbitrary category. However, this is not as simple as it may sound: nowhere in the axioms in the definition of a category does it say that the objects of a category have to *be* sets, so we cannot talk about their elements.

In a sense, it's a miracle that this works at all. We are generalizing definitions from set theory, which are almost always given in terms of elements because elements are the only structure sets have, without ever mentioning the elements of the objects we're talking about. We therefore have to find definitions which we can give purely in terms objects and morphisms between them.

Isomorphism

The concept of an isomorphism exists for many mathematical entities. Two groups can be isomorphic, as can two sets or graphs. It turns out that the concept of isomorphism is best understood as a categorical one.

Definition 15 (isomorphism)**.** Let C be a category, $A, B \in \mathrm{Obj}(\mathrm{C})$. A morphism $f \in \mathrm{Hom}(A,B)$ is said to be an isomorphism if there exists a morphism $g \in \mathrm{Hom}(B,A)$ such that

$$g \circ f = \mathrm{id}_A, \qquad \text{and} \qquad f \circ g = \mathrm{id}_B.$$

$$\begin{tikzcd} A \arrow[loop left, "\mathrm{id}_A"] \arrow[r, shift left, "f"] & B \arrow[l, shift left, "g"] \arrow[loop right, "\mathrm{id}_B"] \end{tikzcd}$$

We usually denote such a g by f^{-1}.

If we have an isomorphism $f\colon A \to B$, we say that A and B are isomorphic, and write $A \simeq B$. This is not an abuse of notation; it is easy to check that isomorphism is an equivalence relation.

Monomorphisms

Monomorphicity is an attempt to define a property analogous to injectivity which can be used in any category. The regular definition of injectivity, i.e.

$$f(a) = f(b) \implies a = b$$

will not do, because the objects in a category do not in general have elements. Therefore, we have to use a different property of injective functions: that they are left-cancellable. That is, if f is injective, then

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

for any functions g_1 and g_2.

Definition 16 (monomorphism)**.** Let C be a category, $A, B \in \mathrm{Obj}(\mathsf{C})$. A morphism $f\colon A \to B$ is said to be a monomorphism (or simply *mono*) if for any $Z \in \mathrm{Obj}(\mathsf{C})$ and any $g_1, g_2\colon Z \to A$, $f \circ g_1 = f \circ g_2$ implies $g_1 = g_2$.

$$\begin{tikzcd} Z \arrow[r, shift left, "g_1"] \arrow[r, shift right, "g_2"'] & A \arrow[r, "f"] & B \end{tikzcd}$$

Note 17. When we wish to notationally distinguish monomorphisms, we will denote them by hooked arrows: if $f\colon A \to B$ is mono, we will write

$$\begin{tikzcd} A \arrow[r, hook, "f"] & B \end{tikzcd} .$$

This turns out to work due to the following theorem.

Theorem 18. In Set, a morphism is a monomorphism if and only if it is injective.

Proof. Suppose $f\colon A \to B$ is a monomorphism. Then for any set Z and any maps $g_1, g_2\colon Z \to A$, $f \circ g_1 = f \circ g_2$ implies $g_1 = g_2$. In particular, take Z to be the singleton $Z = \{*\}$ and call $g_1(*) = a_1$ and $g_2(*) = a_2$. Then $(f \circ g_1)(*) = f(a_1)$ and $(f \circ g_2)(*) = f(a_2)$, so

$$f(a_1) = f(a_2) \implies a_1 = a_2.$$

But this is exactly the definition of injectivity.

Now suppose that f is injective. Then for any Z and g_1, g_2 as above,

$$(f \circ g_1)(z) = (f \circ g_2)(z) \implies g_1(z) = g_2(z) \qquad \text{for all } z \in Z.$$

But this means that $g_1 = g_2$, so f is mono. □

Example 19. In Vect_k, a morphism $L\colon V \to W$ is mono if and only if it is injective. That injectivity implies monomorphicity is obvious because any linear map is in particular a set function. To see that any linear monomorphism is injective, consider k as a vector space over itself. Then for any maps $A, B\colon k \to V$, we have

$$L \circ A = L \circ B \implies A = B.$$

In particular, for any $a, b \in V$, let $A(1) = a$ and $B(1) = b$. Then

$$\begin{aligned} L(a) &= L(b) \\ L(A(1)) &= L(B(1)) \\ (L \circ A)(1) &= (L \circ B)(1) \\ L \circ A &= L \circ B \\ A &= B \\ A(1) &= B(1) \\ a &= b. \end{aligned}$$

Epimorphisms

The notion of a surjection can also be generalized. Pleasingly, it is more clear that epimorphisms are dual to monomorphisms than that surjectivity is dual to injectivity.

Definition 20 (epimorphism)**.** Let C be a category, $A, B \in \text{Obj}(\mathsf{C})$. A morphism $f : A \to B$ is said to be a epimorphism if for all $Z \in \text{Obj}(\mathsf{C})$ and all $g_1, g_2 : B \to Z$, $g_1 \circ f = g_2 \circ f$ implies $g_1 = g_2$.

$$\begin{tikzcd} A \arrow[r, "f"] & B \arrow[r, shift left, "g_1"] \arrow[r, shift right, "g_2"'] & Z \end{tikzcd} .$$

Notation 21. We will denote epimorphisms by two-headed arrows. That is, if $f : A \to B$ is epi, we will write

$$\begin{tikzcd} A \arrow[r, two heads, "f"] & B \end{tikzcd}$$

Theorem 22. In Set, a morphism f is an epimorphism if and only if it is a surjection.

Proof. Suppose $f : A \to B$ is an epimorphism. Then for any set Z and maps $g_1, g_2 : B \to Z$, there $g_1 \circ f = g_2 \circ f$ implies $g_1 = g_2$. In particular, this is true if $Z = \{0, 1\}$, and g_1 and g_2 are as follows.

$$g_1 : b \mapsto 0 \quad \text{for all } b; \qquad g_2 : b \mapsto \begin{cases} 0, & b \in \text{im}(f) \\ 1, & b \notin \text{im}(f). \end{cases}$$

But then $g_1 \circ f = g_2 \circ f$, so $g_1 = g_2$ since f is epi. Hence, $\text{im}(f) = B$, so f is surjective.

If f is surjective, then there exists a right inverse f^{-1} such that $f^{-1} \circ f = \text{id}_A$. Then

$$g_1 \circ f = g_2 \circ f \implies g_1 \circ f \circ f^{-1} = g_2 \circ f \circ f^{-1} \implies g_1 = g_2.$$

□

Example 23. In Vect_k, epimorphisms are surjective linear maps.

Note 24. In Set, we have the following correspondences.

$$\text{Isomorphism} \iff \text{Bijective} \iff \text{Injective \& Surjective} \iff \text{Monic \& Epic}$$

Thus in Set a morphism is an isomorphism if and only if it is a monomorphism and an epimorphism. A very common mistake is to assume that this holds true in any category. This is wrong! A morphism can be monic and epic without being an isomorphism.

This occurrs, for example, in the category Top, whose objects are topological spaces and whose morphisms are continuous maps. This is because of the following chain of reasoning.

- In order for a continuous map f to be monic, it is certainly sufficient that it be injective, since the injectivity of a function implies that for any other functions α and β,

  $$f \circ \alpha = f \circ \beta \implies \alpha = \beta.$$

  This certainly remains true if α and β are taken from the subset of set-functions which are continuous. Thus in Top,

  $$\text{injective} \implies \text{monic}.$$

- Exactly analogous reasoning shows that in Top,

  $$\text{surjective} \implies \text{epic}.$$

- Thus, any continuous map which is bijective is both mono and epic in Top. However, there are continuous, bijective maps which are not isomorphisms! Take, for example, the map

  $$[0, 2\pi) \to S^1 \subset \mathbb{R}^2; \qquad x \mapsto (\cos x, \sin x).$$

  This is clearly continuous and bijective, hence both a monomorphism and an epimorphism in Top. However, it has no inverse in Top: its inverse function is discontinuous at $(1, 0) \in S^1$, so it is not a morphism $S^1 \to [0, 2\pi)$.

Smallness and local smallness

Set theory has some foundational annoyances. Among the most famous of these is Russel's paradox, which demonstrates that that not every definable collection is a set. Category theory has its own foundational issues, which for the most part we will avoid. However, there are a few important situations in which foundational questions of size play an unavoidably important role.

Recall that in our definition of a category, we said that objects and morphisms of a category need not fit inside a set; we only need them to form a 'collection.' We were cagy about what exactly we meant by this. The more precise statement is that they must form a *class,* although we will not have to worry about what a class is.

Using classes instead of sets works perfectly well. However, sets are better behaved than classes, and it is useful to have a special name for categories whose objects and/or morphisms *really do* fit into a set.

Definition 25 (small, locally small, hom-set)**.** A category C can have the following properties.

- We say that C is locally small if for all $A, B \in \mathrm{Obj}(\mathsf{C})$, $\mathrm{Hom}_{\mathsf{C}}(A, B)$ is a small set.
- We say that C is small if $\mathrm{Obj}(\mathsf{C})$ is a small set and for all objects $A, B \in \mathrm{Obj}(\mathsf{C})$, $\mathrm{Hom}_{\mathsf{C}}(A, B)$ is a small set.

If we are working with a category which is locally small, so that $\mathrm{Hom}_{\mathsf{C}}(A, B)$ is always a small set, we call $\mathrm{Hom}_{\mathsf{C}}(A, B)$ the hom-set. (Actually, terminology is often abused, and $\mathrm{Hom}_{\mathsf{C}}(A, B)$ is called a hom-set even if it is not a set.)

Example 26. The category Set is locally small but not small.

1.2 Functors

Category theory is as fantastically useful as it is because of the ubiquity of the notion of a structure preserving map. Many interesting examples of categories arise by considering as objects some class of mathematical entities, and as morphisms the structure-preserving maps between them.

Categories are themselves mathematical objects, so it would be hypocrytical not to look for the correct notion of a structure preserving map between them. This is called a functor.

Definition 27 (functor)**.** Let C and D be categories. A functor $\mathcal{F}$ from C to D is the following.

- It assigns to each object $X \in \mathrm{Obj}(\mathsf{C})$ an object $\mathcal{F}(X) \in \mathrm{Obj}(\mathsf{D})$.
- It assigns each morphism $f \in \mathrm{Hom}_{\mathsf{C}}(X, Y)$ to a morphism $\mathcal{F}(f) \in \mathrm{Hom}(\mathcal{F}(X), \mathcal{F}(Y))$ in such a way that the following conditions are satisfied.
  - It maps identities to identities, i.e.

$$\mathcal{F}(\mathrm{id}_X) = \mathrm{id}_{\mathcal{F}(X)} \qquad \text{for all } X \in \mathrm{Obj}(\mathsf{C}).$$

  - It respects composition:

$$\mathcal{F}(g \circ f) = \mathcal{F}(g) \circ \mathcal{F}(f).$$

Notation 28. We will typeset functors with calligraphic letters, and notate them with squiggly arrows. For example, if C and D are categories and $\mathcal{F}$ is a functor from C to D, then we would write

$$\mathcal{F} \colon \mathsf{C} \rightsquigarrow \mathsf{D}.$$

Later, when we confront the idea that functors are really just morphisms in the category of categories, we will stop using the squiggly-arrow notation and denote functors with straight arrows. For now, however, it will be helpful to keep morphisms straight from functors.

1.2.1 Examples

Example 29. Let $\mathsf{1}$ be the category with one object $*$ (Example 4). Let C be any category. Then for each $X \in \mathrm{Obj}(\mathsf{C})$, we have the functor

$$\mathcal{F}_X \colon \mathsf{1} \rightsquigarrow \mathsf{C}; \qquad \mathcal{F}(*) = X, \quad \mathcal{F}(\mathrm{id}_*) = \mathrm{id}_X.$$

There is a special name for functors whose domain is a product category.

Definition 30 (bifunctor)**.** A bifunctor is a functor whose domain is a product category (Definition 11).

Example 31. The Cartesian product of sets is a bifunctor

$$\times \colon \mathsf{Set} \times \mathsf{Set} \to \mathsf{Set},$$

which sends an object (i.e. a pair of sets) (A, B) to its Cartesian product $A \times B$, and sends a morphism

$$(f, g) \colon (A, B) \to (C, D)$$

to a morphism

$$f \times g \colon A \times B \to C \times D; \qquad (a, b) \mapsto (f(a), g(b)).$$

Checking that this respects composition and sends identities to identities is not at all tricky.

Example 32. Let G be a group, and V a vector space over some field k. A *representation* ρ of G on V assigns to each $g \in G$ a linear transformation $\rho(g) \colon V \to V$ such that

$$\rho(gh) = \rho(g) \cdot \rho(h).$$

Let G be a group. Recall Example 8, in which we constructed from a group G a category BG with one object $*$, whose morphisms $\mathrm{Hom}_{BG}(*, *)$ were given by the elements of G.

Then functors $\rho \colon BG \rightsquigarrow \mathsf{Vect}_k$ are k-linear representations of G!

To see this, let us unwrap the definition. The functor ρ assigns to $* \in \mathrm{Obj}(BG)$ an object $\rho(*) = V \in \mathrm{Obj}(\mathsf{Vect})$, and to each morphism $g \colon * \to *$ a morphism $\rho(g) \colon V \to V$. Furthermore, ρ must respect composition in the sense that

$$\rho(g \cdot h) = \rho(g) \cdot \rho(h).$$

But this means precisely that ρ is a representation of G.

Example 33. Recall that Grp is the category of groups. Denote by CRing the category of commutative rings.

The following are functors $\mathsf{CRing} \rightsquigarrow \mathsf{Grp}$.

- GL_n, which assigns to each commutative ring K the group of all $n \times n$ invertible matrices with entries in K, and to each commutative ring homomorphism $f \colon K \to K'$ a map
$$\mathrm{GL}_n(f) \colon \mathrm{GL}_n(K) \to \mathrm{GL}_n(K'); \qquad \begin{pmatrix} a_{11} & \cdots & a_{nn} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \mapsto \begin{pmatrix} f(a_{11}) & \cdots & f(a_{nn}) \\ \vdots & \ddots & \vdots \\ f(a_{n1}) & \cdots & f(a_{nn}) \end{pmatrix}.$$
- $(\,\cdot\,)^*$, which maps each commutative ring K to its group of units K^*, and each morphism $K \to K'$ to its restriction to K^*.

Example 34. Let $\mathcal{P}(\{0, 1\})$ be the poset category from Example 5. A functor $\mathcal{P}(\{0, 1\}) \rightsquigarrow \mathsf{C}$ is a commuting square in C. Similarly, a functor $\mathcal{P}(\{0, 1, 2\}) \rightsquigarrow \mathcal{C}$ is a cube whose faces commute.

Example 35. Let k be a field. There is a functor $(-)^* \colon \mathsf{Vect}_k^{\mathrm{op}} \rightsquigarrow \mathsf{Vect}_k$ which sends each vector space to its dual.

1.2.2 Properties of functors

Full, faithful, and essentially surjective

Definition 36 (full, faithful)**.** Let C and D be locally small categories (Definition 25), and let $\mathcal{F} : \mathsf{C} \rightsquigarrow \mathsf{D}$. Then $\mathcal{F}$ induces a family of set-functions

$$\mathcal{F}_{X,Y} \colon \mathrm{Hom}_{\mathsf{C}}(X, Y) \to \mathrm{Hom}_{\mathsf{D}}(\mathcal{F}(X), \mathcal{F}(Y)).$$

We say that $\mathcal{F}$ is

- full if $\mathcal{F}_{X,Y}$ is surjective for all $X, Y \in \mathrm{Obj}(\mathsf{C})$
- faithful if $\mathcal{F}_{X,Y}$ is injective for all $X, Y \in \mathrm{Obj}(\mathsf{C})$,
- fully faithful if $\mathcal{F}$ is full and faithful.

Note 37. Fullness and faithfulness are *not* the functorial analogs of surjectivity and injectivity. A functor between small categories can be full (resp. faithful) without being surjective (resp. injective) on objects. Instead, we have the following result.

Lemma 38. A fully faithful functor is injective on objects up to isomorphism. That is, if $\mathcal{F}\colon \mathsf{C} \rightsquigarrow \mathsf{D}$ is a fully faithful functor and $\mathcal{F}(A) \simeq \mathcal{F}(B)$, then $A \simeq B$.

Proof. Let $\mathcal{F}\colon \mathsf{C} \rightsquigarrow \mathsf{D}$ be a fully faithful functor, and suppose that $\mathcal{F}(A) \simeq \mathcal{F}(B)$. Then there exist $f'\colon \mathcal{F}(A) \to \mathcal{F}(B)$ and $g'\colon \mathcal{F}(B) \to \mathcal{F}(A)$ such that $f' \circ g' = \mathrm{id}_{\mathcal{F}(B)}$ and $g' \circ f' = \mathrm{id}_{\mathcal{F}(A)}$. Because the function $\mathcal{F}_{A,B}$ is bijective it is invertible, so there is a unique morphism $f \in \mathrm{Hom}_{\mathsf{C}}(A, B)$ such that $\mathcal{F}(f) = f'$, and similarly there is a unique $g \in \mathrm{Hom}_{\mathsf{C}}(B, A)$ such that $\mathcal{F}(g) = g'$.

Now,

$$\mathrm{id}_{\mathcal{F}(A)} = g' \circ f' = \mathcal{F}(g) \circ \mathcal{F}(f) = \mathcal{F}(g \circ f),$$

and since $\mathcal{F}$ is injective, we must have $g \circ f = \mathrm{id}_A$. Identical logic shows that we must also have $f \circ g = \mathrm{id}_B$. Thus $A \simeq B$. ☐

It now makes sense to define a different functorial version of surjectivity.

Definition 39 (essentially surjective)**.** A functor $\mathcal{F}\colon \mathsf{C} \rightsquigarrow \mathsf{D}$ is essentially surjective if for every $A' \in \mathrm{Obj}(\mathsf{D})$, there exists $A \in \mathrm{Obj}(\mathsf{C})$ such that $A' \simeq \mathcal{F}(A)$.

Covariance and contravariance

What we have called simply a *functor*, many people would call a *covariant functor.* These people would say that there is a second kind of functor, called a *contravariant functor,* which flips arrows around. That is, applying a contravariant functor $\mathcal{F}$ to a morphism $f\colon A \to B$, one would find a morphism

$$\mathcal{F}(f)\colon \mathcal{F}(B) \to \mathcal{F}(A).$$

In order to respect compositions, these people say, a contravariant functor must obey the modified composition rule

$$\mathcal{F}(f \circ g) = \mathcal{F}(g) \circ \mathcal{F}(f).$$

We are choosing not to use the notion of a contravariant functor, and for a good reason: we don't need it! Recall the so-called opposite category (Definition 10). Roughly speaking, the opposite to a category is a category in which arrows go the other way.

A contravariant functor $\mathcal{F}\colon \mathsf{C} \rightsquigarrow \mathsf{D}$, then, can be thought of as a *covariant* (i.e. ordinary) functor $\mathcal{F}\colon \mathsf{C}^{\mathrm{op}} \rightsquigarrow\!\!\longrightarrow \mathcal{D}$. This turns out to be much simpler and more convenient than dealing with two kinds of functors all the time.

Definition 40 (contravariant functor)**.** A contravariant functor $\mathcal{F}\colon \mathsf{C} \rightsquigarrow \mathsf{D}$ is simply a (covariant) functor $\mathcal{F}\mathsf{C}^{\mathrm{op}} \rightsquigarrow \mathsf{D}$.

1.3 Natural transformations

Saunders Mac Lane, one of the fathers of category theory, used to say that he invented categories so he could talk about functors, and he invented functors so he could talk about natural transformations. Indeed, arguably the first paper ever published on category theory, published by Eilenberg and Mac Lane in 1945 and titled “General Theory of Natural Equivalences” [?], was about natural transformations.

Natural transformation can be thought of at several different levels of abstraction. At their most abstract, they provide a notion of ‘morphisms between functors,’ and indeed we will later see that natural transformations provide the correct notion of morphisms for a category whose objects are functors. However, they were originally studied not for this reason, but because they provide a rigorous footing for the notion of canonicalness itself.

For example, one often hears the statement “The cartesian product of sets is commutative (up to isomorphism) because for any two sets A and B, $A{\times}B$ is isomorphic to $B{\times}A$.” This is certainly true, but not very meaningful; any two sets with the same cardinality are isomorphic. What one is trying to say is that $A \times B$ and $B \times A$ are *canonically* isomorphic. The question is: what exactly is meant by ‘canonically’?

This is a surprisingly elusive question. For any two sets A and B, the canonical isomorphism in question is the obvious map

$$\mathrm{swap}_{A,B} \colon A \times B \to B \times A; \qquad (a, b) \mapsto (b, a).$$

However, picking some other sets C and D gives is a different swap function. In what sense are these swap functions instances of the same ‘swap operation?’ And what does this have to do with canonicalness?

One way of formalizing the sameness comes from the fact that this swap function commutes with functional evaluation. Notice that for *any* functions $f \colon A \to C$ and $g \colon B \to D$, the following diagram commutes.

$$\begin{tikzcd}
A \times B \arrow[rrr, "{(f,g)}"] \arrow[ddd, "\mathrm{swap}_{A,B}"'] & & & C \times D \arrow[ddd, "\mathrm{swap}_{C,D}"] \\
& (a,b) \arrow[r, mapsto] \arrow[d, mapsto] & (f(a),g(b)) \arrow[d, mapsto] & \\
& (b,a) \arrow[r, mapsto] & (g(b),f(a)) & \\
B \times A \arrow[rrr, "{(g,f)}"'] & & & D \times C
\end{tikzcd}$$

That is to say, it doesn’t matter whether you use first use f and g to map $A \times B$ into $C \times D$ and then use the swap isomorphism, or whether you do the swap first and then use f and g to map $B \times A$ into $D \times C$; you get the same map either way. Furthermore, this is true for *any* maps f and g.

As we have seen (in Example 31), the cartesian product of sets is a functor from the product category (Definition 11)

$$\times \colon \mathsf{Set} \times \mathsf{Set} \rightsquigarrow \mathsf{Set}; \qquad (A, B) \mapsto A \times B.$$

There is another 'twisted' Cartesian product functor

$$\tilde{\times} \colon \mathsf{Set} \times \mathsf{Set} \rightsquigarrow \mathsf{Set}; \qquad (A, B) \mapsto B \times A.$$

As we will see in the rest of this chapter, the above discussion means precisely that the functions $\mathrm{swap}_{A,B}$ form a natural transformation, called swap, between $\times$ and $\tilde{\times}$.

This example makes clear a common theme of natural transformations: they formalize the idea of 'the same function between different objects.' They allow us to make precise statements like '$\mathrm{swap}_{A,B}$ and $\mathrm{swap}_{C,D}$ are somehow the same, despite the fact that they cannot be equal as functions.'

Definitions and elementary examples

Definition 41 (natural transformation)**.** let C and D be categories, and let $\mathcal{F}$ and $\mathcal{G}$ be functors from C to D. A natural transformation η between $\mathcal{F}$ and $\mathcal{G}$ consists of

- for each object $A \in \mathrm{Obj}(\mathsf{C})$ a morphism $\eta_A \colon \mathcal{F}(A) \to \mathcal{G}(A)$, such that
- for all $A, B \in \mathrm{Obj}(C)$, for each morphism $f \in \mathrm{Hom}(A, B)$, the diagram

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

commutes.

We will use double-shafted arrows to denote natural transformations: if $\mathcal{F}$ and $\mathcal{G}$ are functors and η is a natural transformation from $\mathcal{F}$ to $\mathcal{G}$, we will write

$$\eta \colon \mathcal{F} \Rightarrow \mathcal{G}.$$

Definition 42 (natural isomorphism)**.** A natural isomorphism $\eta \colon \mathcal{F} \Rightarrow \mathcal{G}$ is a natural transformation such that each η_A is an isomorphism.

Lemma 43. Given a natural isomorphism $\eta \colon \mathcal{F} \Rightarrow \mathcal{G}$, we can construct an inverse natural isomorphism $\eta^{-1} \colon \mathcal{G} \Rightarrow \mathcal{F}$.

Proof. The natural transformation gives us for any two objects A and B and morphism $f \colon A \to B$ a naturality square

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

which tells us that

$$\eta_B \circ \mathcal{F}(f) = \mathcal{G}(f) \circ \eta_A.$$

Since η is a natural isomorphism, its components η_A are isomorphisms, so they have inverses η_A^{-1}. Acting on the above equation with η_A^{-1} from the right and η_B^{-1} from the left, we find

$$\mathcal{F}(f) \circ \eta_A^{-1} = \eta_B^{-1} \circ \mathcal{G}(f),$$

i.e. the following diagram commutes.

$$\begin{tikzcd}
\mathcal{G}(A) \arrow[r, "\mathcal{G}(f)"] \arrow[d, "\eta_A^{-1}"'] & \mathcal{G}(B) \arrow[d, "\eta_B^{-1}"] \\
\mathcal{F}(A) \arrow[r, "\mathcal{F}(f)"] & \mathcal{F}(B)
\end{tikzcd}$$

But this is just the naturality square for a natural isomorphism η^{-1} with components $(\eta^{-1})_A = \eta_A^{-1}$. ☐

Example 44 (the determinant). Recall the functors GL_n and $(\cdot)^*$ from Example 33.

Denote by $\det_K M$ the determinant of a matrix M with its entries in a commutative ring K. Then the determinant is a map

$$\det_K \colon \mathrm{GL}_n(K) \to K^*.$$

Because the determinant is defined by the same formula for each K, the action of f commutes with $\det_K$: it doesn't matter whether we map the entries of M with f first and then take the determinant, or take the determinant first and then feed the result to f.

That is to say, the following diagram commutes.

$$\begin{tikzcd}
\mathrm{GL}_n(K) \arrow[r, "\det_K"] \arrow[d, "\mathrm{GL}_n(f)"'] & K^* \arrow[d, "f^*"] \\
\mathrm{GL}_n(K') \arrow[r, "\det_{K'}"] & K'^*
\end{tikzcd}$$

This means that det is a natural transformation $\mathrm{GL}_n \Rightarrow (\,\cdot\,)^*$.

Example 45 (intertwiners and scattering theory). Let G be a group, BG its delooping groupoid, and let ρ and $\rho' \colon \mathrm{G} \rightsquigarrow \mathrm{Vect}$ be representations (recall Example 32). Let $\rho(*) = V$ and $\rho'(*) = W$.

A natural transformation $\eta \colon \rho \Rightarrow \rho'$ is called an *intertwiner*. So what *is* an intertwiner?

The natural transformation η has only one component, $\eta_* \colon V \to W$, which is subject to the condition that for any $g \in \mathrm{Hom}_{BG}(*, *)$, the diagram below commutes.

$$\begin{tikzcd}
V \arrow[r, "\rho(g)"] \arrow[d, "\eta_*"'] & V \arrow[d, "\eta_*"] \\
W \arrow[r, "\rho'(g)"] & W
\end{tikzcd}$$

That is, an intertwiner is a linear map $V \to W$ such that for all g,

$$\eta_* \circ \rho(g) = \rho'(g) \circ \eta_*.$$

Intertwiners are extremely important in high-energy physics.

In scattering theory one models the state of the world as $T \to -\infty$ as a composite system of several incoming free particles, and as $T \to +\infty$ a composite system of outgoing free particles. Therefore, the Hilbert space of the incoming particles is

$$\mathscr{H}_{\text{in}} = \mathscr{H}_{1,\text{in}} \otimes \cdots \otimes \mathscr{H}_{m,\text{in}}$$

and that of the outgoing particles is

$$\mathscr{H}_{\text{out}} = \mathscr{H}_{1,\text{out}} \otimes \cdots \otimes \mathscr{H}_{n,\text{out}}.$$

In relativistic quantum mechanics, the Hilbert space of a free particle alone in the world is the representation space of some irreducible representation of the Poincaré group. The action of the Poincaré group on this representation space implements Poincaré transformations on the particle. For a system of several particles, the Poincaré group acts on the incoming and outgoing Hilbert spaces via the tensor product representation.

Scattering theory finds a map $U\colon \mathscr{H}_{\text{in}} \to \mathscr{H}_{\text{out}}$, which interpolates between incoming states and outgoing states. However, not any linear map will do; since we demand that our laws be Poincaré-invariant, we want our transformation to be Poincaré-equivariant. Therefore, we want scattering theory to produce for us a map $U\colon \mathscr{H}_{\text{in}} \to \mathscr{H}_{\text{out}}$ such that the diagram

$$\begin{tikzcd}
\mathscr{H}_{\text{in}} \arrow[r, "\rho_{\text{in}}(g)"] \arrow[d, "U"'] & \mathscr{H}_{\text{in}} \arrow[d, "U"] \\
\mathscr{H}_{\text{out}} \arrow[r, "\rho_{\text{out}}(g)"] & \mathscr{H}_{\text{out}}
\end{tikzcd}$$

commutes.

That is, scattering theory produces intertwiners between the incoming and outgoing representations ρ_{in} and ρ_{out} of the Poincaré group.

Vertical composition

We mentioned earlier that one could define a category whose objects were functors and whose morphisms were natural transformations. We are now going to make good on that threat. In order to do so, we need only specify how to compose natural transformations. The correct notion is known as *vertical composition.*

Lemma 46. Let C and D be categories, $\mathcal{F}$, $\mathcal{G}$, $\mathcal{H}$ be functors, and Φ and Ψ be natural transformations as follows.

$$\begin{tikzcd}
\mathsf{C} \arrow[r, rightsquigarrow, bend left=60, "\mathcal{F}"{name=F}] \arrow[r, rightsquigarrow, "\mathcal{G}"{name=G}] \arrow[r, rightsquigarrow, bend right=60, "\mathcal{H}"'{name=H}] & \mathsf{D}
\arrow[Rightarrow, from=F, to=G, "\Phi"]
\arrow[Rightarrow, from=G, to=H, "\Psi"]
\end{tikzcd}$$

This induces a natural transformation $\mathcal{F} \Rightarrow \mathcal{H}$.

Proof. For each object $A \in \text{Obj}(\mathsf{C})$, the composition $\Psi_A \circ \Phi_A$ exists and maps $\mathcal{F}(A) \to \mathcal{H}(A)$. Let's write

$$\Psi_A \circ \Phi_A = (\Psi \circ \Phi)_A.$$

We have to show that these are the components of a natural transformation, i.e. that they make the following diagram commute for all $A, B \in \mathrm{Obj}(\mathsf{C})$, all $f : A \to B$.

$$\begin{tikzcd}
\mathcal{F}(A) \arrow[r, "\mathcal{F}(f)"] \arrow[d, "(\Psi\circ\Phi)_A"'] & \mathcal{F}(\mathcal{B}) \arrow[d, "(\Psi\circ\Phi)_B"] \\
\mathcal{H}(A) \arrow[r, "\mathcal{H}(f)"] & \mathcal{H}(B)
\end{tikzcd}$$

We can do this by adding a middle row.

$$\begin{tikzcd}
\mathcal{F}(A) \arrow[r, "\mathcal{F}(f)"] \arrow[d, "\Phi_A"'] & \mathcal{F}(\mathcal{B}) \arrow[d, "\Phi_B"] \\
\mathcal{G}(A) \arrow[r, "\mathcal{G}(f)"] \arrow[d, "\Psi_A"'] & \mathcal{G}(B) \arrow[d, "\Psi_B"] \\
\mathcal{H}(A) \arrow[r, "\mathcal{H}(f)"] & \mathcal{H}(B)
\end{tikzcd}$$

The top and bottom squares are the naturality squares for Φ and Ψ respectively. The outside square is the one we want to commute, and it manifestly does because each of the inside squares does. □

Definition 47 (vertical composition). The above composition $\Psi \circ \Phi$ is called vertical composition.

One of the lessons of category theory is that any time we see a collection of things and mappings between them, turning them into a category will be a good investment. We saw that functors were mappings between categories, and defined a category Cat whose objects are categories and whose morphisms are functors. We can now define a category whose objects are functors and whose morphisms are natural transformations.

Definition 48 (functor category). Let C and D be categories. The functor category $\mathsf{Fun}(\mathsf{C}, \mathsf{D})$ (sometimes D^{C} or $[\mathsf{C}, \mathsf{D}]$) is the category whose objects are functors $\mathsf{C} \rightsquigarrow \mathsf{D}$, and whose morphisms are natural transformations between them. The composition is given by vertical composition.

Horizontal composition

We can also compose natural transformations in a not so obvious way.

Lemma 49. Consider the following arrangement of categories, functors, and natural transformations.

$$\begin{tikzcd}
\mathsf{C} \arrow[r, bend left=50, squiggly, "\mathcal{F}"{name=U1}] \arrow[r, bend right=50, squiggly, "\mathcal{F}'"'{name=D1}] & \mathsf{D} \arrow[r, bend left=50, squiggly, "\mathcal{G}"{name=U2}] \arrow[r, bend right=50, squiggly, "\mathcal{G}'"'{name=D2}] & \mathsf{E}
\arrow[Rightarrow, from=U1, to=D1, "\Phi"]
\arrow[Rightarrow, from=U2, to=D2, "\Psi"]
\end{tikzcd}$$

This induces a natural transformation $\mathcal{G} \circ \mathcal{F} \Rightarrow \mathcal{G}' \circ \mathcal{F}'$.

Proof. By definition, Φ and Ψ make the diagrams

$$\begin{tikzcd}
\mathcal{F}(A) \arrow[r, "\mathcal{F}(f)"] \arrow[d, swap, "\Phi_A"] & \mathcal{F}(B) \arrow[d, "\Phi_B"] \\
\mathcal{F}'(A) \arrow[r, swap, "\mathcal{F}'(f)"] & \mathcal{F}'(B)
\end{tikzcd}
\quad \text{and} \quad
\begin{tikzcd}
\mathcal{G}(A) \arrow[r, "\mathcal{G}(f)"] \arrow[d, swap, "\Psi_A"] & \mathcal{G}(B) \arrow[d, "\Psi_B"] \\
\mathcal{G}'(A) \arrow[r, swap, "\mathcal{G}'(f)"] & \mathcal{G}'(B)
\end{tikzcd}$$

commute. Since functors respect composition they take commutative diagrams to commutative diagrams, so if we map everything in the first diagram to E with $\mathcal{G}$ to get another commutative diagram in E.

$$\begin{tikzcd}
(\mathcal{G} \circ \mathcal{F})(A) \arrow[r, "(\mathcal{G}\circ\mathcal{F})(f)"] \arrow[d, swap, "\mathcal{G}(\Phi_A)"] & (\mathcal{G} \circ \mathcal{F})(B) \arrow[d, "\mathcal{G}(\Phi_B)"] \\
(\mathcal{G} \circ \mathcal{F}')(A) \arrow[r, "(\mathcal{G}\circ\mathcal{F}')(f)"] & (\mathcal{G} \circ \mathcal{F}')(B)
\end{tikzcd}$$

Since $\mathcal{F}'(f)\colon \mathcal{F}'(A) \longrightarrow \mathcal{F}'(B)$ is a perfectly good example of a morphism in D, the natural transformation and Ψ has components $\Psi_{\mathcal{F}'(A)}$ and $\Psi_{\mathcal{F}'(B)}$ which makes the following diagram commutes.

$$\begin{tikzcd}
(\mathcal{G} \circ \mathcal{F}')(A) \arrow[r, "(\mathcal{G}\circ\mathcal{F}')(f)"] \arrow[d, swap, "\Psi_{\mathcal{F}'(A)}"] & (\mathcal{G} \circ \mathcal{F}')(B) \arrow[d, "\Psi_{\mathcal{F}'(B)}"] \\
(\mathcal{G}' \circ \mathcal{F}')(A) \arrow[r, "(\mathcal{G}'\circ\mathcal{F}')(f)"] & (\mathcal{G}' \circ \mathcal{F}')(B)
\end{tikzcd}$$

Sticking these two diagrams on top of each other gives a new commutative diagram.

$$\begin{tikzcd}
(\mathcal{G} \circ \mathcal{F})(A) \arrow[r, "(\mathcal{G}\circ\mathcal{F})(f)"] \arrow[d, swap, "\mathcal{G}(\Phi_A)"] & (\mathcal{G} \circ \mathcal{F})(B) \arrow[d, "\mathcal{G}(\Phi_B)"] \\
(\mathcal{G} \circ \mathcal{F}')(A) \arrow[r, "(\mathcal{G}\circ\mathcal{F}')(f)"] \arrow[d, swap, "\Psi_{\mathcal{F}'(A)}"] & (\mathcal{G} \circ \mathcal{F}')(B) \arrow[d, "\Psi_{\mathcal{F}'(B)}"] \\
(\mathcal{G}' \circ \mathcal{F}')(A) \arrow[r, "(\mathcal{G}'\circ\mathcal{F}')(f)"] & (\mathcal{G}' \circ \mathcal{F}')(B)
\end{tikzcd}$$

The outside rectangle is nothing else but the commuting square for a natural transformation

$$(\Psi * \Phi)\colon \mathcal{G} \circ \mathcal{F} \Rightarrow \mathcal{G}' \circ \mathcal{F}'$$

with components $(\Psi * \Phi)_A = \Psi_{\mathcal{F}(A)} \circ \mathcal{G}(\Phi_A)$. □

Definition 50 (horizontal composition)**.** The natural transformation $\Psi * \Phi$ defined above by by

$$(\Psi * \Phi)_A = \Psi_{\mathcal{F}(A)} \circ G(\Phi(A))$$

is called the horizontal composition of Φ and Ψ.

One might worry the above definition of the horizontal composition is lopsided. Why did we apply the functor $\mathcal{G}$ to the first square rather than $\mathcal{G}'$? We find the following fact reassuring.

If we had applied the functor $\mathcal{G}'$ instead of $\mathcal{G}$, we would have found the following.

$$\begin{tikzcd}
(\mathcal{G}' \circ \mathcal{F})(A) \arrow[r, "(\mathcal{G}' \circ \mathcal{F})(f)"] \arrow[d, "\mathcal{G}'(\Phi_A)"'] & (\mathcal{G}' \circ \mathcal{F})(B) \arrow[d, "\mathcal{G}'(\Phi_B)"] \\
(\mathcal{G}' \circ \mathcal{F}')(A) \arrow[r, "(\mathcal{G}' \circ \mathcal{F}')(f)"] & (\mathcal{G}' \circ \mathcal{F}')(B)
\end{tikzcd}$$

Then we could have glued to it the bottom of the following commuting square.

$$\begin{tikzcd}
(\mathcal{G} \circ \mathcal{F})(A) \arrow[r, "(\mathcal{G} \circ \mathcal{F})(f)"] \arrow[d, "\Psi_{\mathcal{F}(A)}"'] & (\mathcal{G} \circ \mathcal{F})(B) \arrow[d, "\Psi_{\mathcal{F}(B)}"] \\
(\mathcal{G}' \circ \mathcal{F})(A) \arrow[r, "(\mathcal{G}' \circ \mathcal{F})(f)"] & (\mathcal{G}' \circ \mathcal{F})(B)
\end{tikzcd}$$

If you do this you get *another* natural transformation $\mathcal{G} \circ \mathcal{F} \Rightarrow \mathcal{G}' \circ \mathcal{F}'$, with components $\mathcal{G}'(\Phi_A) \circ \Psi_{\mathcal{F}(A)}$. Why did we use the first definition rather than this one?

It turns out that $\Psi_{\mathcal{F}(A)} \circ \mathcal{G}(\Phi_A)$ and $\mathcal{G}'(\Phi_A) \circ \Psi_{\mathcal{F}(A)}$ are equal. To see this, pick any $A \in \mathrm{Obj}(\mathsf{A})$. From the morphism

$$\Phi_A \colon \mathcal{F}(A) \to \mathcal{F}'(A),$$

the natural transformation Ψ gives us a commuting square

$$\begin{tikzcd}
(\mathcal{G} \circ \mathcal{F})(A) \arrow[r, "\mathcal{G}(\Phi_A)"] \arrow[d, "\Psi_{\mathcal{F}(A)}"'] & (\mathcal{G} \circ \mathcal{F}')(A) \arrow[d, "\Psi_{\mathcal{F}'(A)}"] \\
(\mathcal{G}' \circ \mathcal{F})(A) \arrow[r, "\mathcal{G}'(\Phi_A)"] & (\mathcal{G}' \circ \mathcal{F}')(A)
\end{tikzcd} ;$$

the two ways of going from top left to bottom right are nothing else but $\Psi_{\mathcal{F}(A)} \circ \mathcal{G}(\Phi_A)$ and $\mathcal{G}'(\Phi_A) \circ \Psi_{\mathcal{F}(A)}$.

It doesn't matter whether we first compose horizontally or vertically.

Theorem 51. Suppose we have categories, functors, and natural transformations as follows.

$$\begin{tikzcd}[column sep=huge]
\mathsf{C} \arrow[r, squiggly, bend left=60, "\mathcal{F}", ""{name=F, below}] \arrow[r, squiggly, "\mathcal{G}" description, ""{name=G}] \arrow[r, squiggly, bend right=60, "\mathcal{H}"', ""{name=H, above}] & \mathsf{D} \arrow[r, squiggly, bend left=60, "\mathcal{F}'", ""{name=F2, below}] \arrow[r, squiggly, "\mathcal{G}'" description, ""{name=G2}] \arrow[r, squiggly, bend right=60, "\mathcal{H}'"', ""{name=H2, above}] & \mathsf{E}
\arrow[Rightarrow, from=F, to=G, "\Phi"] \arrow[Rightarrow, from=G, to=H, "\Psi"] \arrow[Rightarrow, from=F2, to=G2, "\Phi'"] \arrow[Rightarrow, from=G2, to=H2, "\Psi'"]
\end{tikzcd}$$

Then

$$(\Psi' \circ \Phi') * (\Psi \circ \Phi) \simeq (\Psi' * \Psi) \circ (\Phi' * \Phi).$$

Proof. We have

$$[(\Psi' \circ \Phi') * (\Psi \circ \Phi)]_c =$$

$\square$

Whiskering

We will mainly be interested in a special case of horizontal composition, in which one of the natural transformations is the identity natural transformation:

Example 52 (whiskering)**.** Consider the following assemblage of categories, functors, and natural transformations.

$$\begin{tikzcd}
\mathsf{C} \arrow[r, squiggly, bend left=50, "\mathcal{F}"{name=U}] \arrow[r, squiggly, bend right=50, "\mathcal{F}'"'{name=L}] & \mathsf{D} \arrow[r, squiggly, "\mathcal{G}"] & \mathsf{E}
\arrow[Rightarrow, from=U, to=L, "\Phi"]
\end{tikzcd}$$

The horizontal composition allows us to Φ to a natural transformation $\mathcal{G} \circ \mathcal{F}$ to $\mathcal{G} \circ \mathcal{F}'$ as follows. First, augment the diagram as follows.

$$\begin{tikzcd}
\mathsf{C} \arrow[r, squiggly, bend left=50, "\mathcal{F}"{name=U}] \arrow[r, squiggly, bend right=50, "\mathcal{F}'"'{name=L}] & \mathsf{D} \arrow[r, squiggly, bend left=50, "\mathcal{G}"{name=U2}] \arrow[r, squiggly, bend right=50, "\mathcal{G}"'{name=L2}] & \mathsf{E}
\arrow[Rightarrow, from=U, to=L, "\Phi"]
\arrow[Rightarrow, from=U2, to=L2, "\mathrm{id}_{\mathcal{G}}"]
\end{tikzcd}$$

We can then take the horizontal composition of Φ and $\mathrm{id}_{\mathcal{G}}$ to get a natural transformation from $\mathcal{G} \circ \mathcal{F}$ to $\mathcal{G} \circ \mathcal{F}'$ with components

$$(\mathrm{id}_{\mathcal{G}} * \Phi)_A = \mathcal{G}(\Phi_A).$$

This natural transformation is called the *right whiskering* of Φ with $\mathcal{G}$, and is denoted $\mathcal{G}\Phi$. That is to say, $(\mathcal{G}\Phi)_A = \mathcal{G}(\Phi_A)$.

$$\begin{tikzcd}
\mathsf{C} \arrow[r, squiggly, bend left=50, "\mathcal{G}\circ\mathcal{F}"{name=U}] \arrow[r, squiggly, bend right=50, "\mathcal{G}\circ\mathcal{F}'"'{name=L}] & \mathsf{E}
\arrow[Rightarrow, from=U, to=L, "\mathcal{G}\Phi"]
\end{tikzcd}$$

The reason for the name is clear: we removed a whisker from the RHS of our diagram.

We can also remove a whisker from the LHS. Given this:

$$\begin{tikzcd}
\mathsf{C} \arrow[r, squiggly, "\mathcal{F}"] & \mathsf{D} \arrow[r, squiggly, bend left=50, "\mathcal{G}"{name=U}] \arrow[r, squiggly, bend right=50, "\mathcal{G}'"'{name=L}] & \mathsf{E}
\arrow[Rightarrow, from=U, to=L, "\Psi"]
\end{tikzcd}$$

we can build a natural transformation (denoted $\Psi\mathcal{F}$) with components

$$(\Psi\mathcal{F})_A = \Psi_{\mathcal{F}(A)},$$

making this:

$$\begin{tikzcd}
\mathsf{C} \arrow[r, squiggly, bend left=50, "\mathcal{G}\circ\mathcal{F}"{name=U}] \arrow[r, squiggly, bend right=50, "\mathcal{G}'\circ\mathcal{F}"'{name=L}] & \mathsf{E}
\arrow[Rightarrow, from=U, to=L, "\Psi\mathcal{F}"]
\end{tikzcd} .$$

This is called the *left whiskering* of Ψ with $\mathcal{F}$

Categorical equivalence

We would like to be able to express when two categories are the ‘same.’ One would think that this would be provided by the notion of an isomorphism, i.e. that two categories C and D should be thought of as the same if there are functors $\mathcal{F}$ and $\mathcal{G}$ between them such that $\mathcal{F} \circ \mathcal{G} = \mathrm{id}_{\mathsf{D}}$ and $\mathcal{G} \circ \mathcal{F} = \mathrm{id}_{\mathsf{C}}$.

$$\begin{tikzcd} \mathsf{C} \arrow[r, squiggly, bend left, "\mathcal{F}"] & \mathsf{D} \arrow[l, squiggly, bend left, "\mathcal{G}"] \end{tikzcd}$$

between them.

However, this is not really in the spirit of category theory. We have seen that functors from one category to another assemble themselves into a category whose morphisms are natural transformations. Demanding $\mathcal{F} \circ \mathcal{G} = \mathrm{id}_{\mathsf{D}}$ would be demanding the equality of objects in a category, which is not usually the right thing to do; we should never demand that two objects be equal, when some notion of isomorphism will do. Therefore, we should demand not that the composition of $\mathcal{F}$ and $\mathcal{G}$ be equal to the identity functor ‘on the nose.’ Instead, we should ask that they be isomorphic, in the sense that there should exist a natural isomorphism between them.

This gives us the following definition.

Definition 53 (categorical equivalence)**.** Let C and D be categories. We say that C and D are equivalent if there is a pair of functors

$$\begin{tikzcd} \mathsf{C} \arrow[r, squiggly, bend left, "\mathcal{F}"] & \mathsf{D} \arrow[l, squiggly, bend left, "\mathcal{G}"] \end{tikzcd}$$

and natural isomorphisms $\eta\colon \mathcal{F} \circ \mathcal{G} \Rightarrow \mathrm{id}_{\mathsf{C}}$ and $\epsilon\colon \mathrm{id}_{\mathsf{C}} \Rightarrow \mathcal{G} \circ \mathcal{F}$.

Note 54. The above definition of categorical equivalence is equivalent to the following: C and D are equivalent if there is a functor $\mathcal{F} : \mathsf{C} \rightsquigarrow \mathsf{D}$ which is fully faithful (Definition 36) and essentially surjective (Definition 39). That is to say, if the equivalence $\mathcal{F}$ is ‘bijective up to isomorphism.’

Definition 55 (essentially small category)**.** A category C is said to be essentially small if it is equivalent to a small category (Definition 25).